QUEST HIERARCHY
FOR
HYPERSPECTRAL FACE RECOGNITION

Dissertation

David M. Ryer, Lt Col, USAF

AFIT/DS/ENS/10-03M

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
Air Force Institute of Technology

Wright-Patterson Air Force Base, Ohio

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT/DS/ENS/10-03M

QUEST HIERARCHY FOR HYPERSPECTRAL FACE RECOGNITION

DISSERTATION

Presented to the Faculty

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

David M. Ryer, BS EE, MS OR

Lieutenant Colonel, USAF

February 2011

AFIT/DS/ENS/10-03M

QUEST HIERARCHY FOR HYPERSPECTRAL FACE RECOGNITION

David M. Ryer, BS EE, MS OR
Lieutenant Colonel, USAF

Approved: Date

_____//SIGNED//_____ 21 MARCH 2011
Dr. Kenneth W. Bauer
Committee Chairman

_____//SIGNED//_____ 21 MARCH 2011
Dr. Michael J. Havrilla
Dean's Representative

_____//SIGNED//_____ 21 MARCH 2011
Dr. Mark E. Oxley
Committee Member

//SIGNED//_____ 21 MARCH 2011
Dr. Steven K. Rogers
Committee Member

Accepted:

____//SIGNED//_____ 21 MARCH 2011
Dr. Marlin U. Thomas
Dean, Graduate School of Engineering and Management

Abstract

Face recognition is an attractive biometric due to the ease in which photographs of the human face can be acquired and processed. The non-intrusive ability of many surveillance systems permits face recognition applications to be used in a myriad of environments. Despite decades of impressive research in this area, face recognition still struggles with variations in illumination, pose and expression not to mention the larger challenge of willful circumvention. The integration of supporting contextual information in a fusion hierarchy known as QUalia Exploitation of Sensor Technology (QUEST) is a novel approach for hyperspectral face recognition that results in performance advantages and a robustness not seen in leading face recognition methodologies.

This research demonstrates a method for the exploitation of hyperspectral imagery and the intelligent processing of contextual layers of spatial, spectral, and temporal information. This approach illustrates the benefit of integrating spatial and spectral domains of imagery for the automatic extraction and integration of novel soft features (biometric). The establishment of the QUEST methodology for face recognition results in an engineering advantage in both performance and efficiency compared to leading and classical face recognition techniques. An interactive environment for the testing and expansion of this recognition framework is also provided.

Acknowledgements

Thank you God for the guidance and support that you give me every day. Thank you to my loving wife Rita, who always encourages me to chase my dreams and then dedicates her life to making them come true. Thank you to my son Shaun who shows me how new and wonderful every day is and for making me the proudest father imaginable. Thank you to my Mom and Dad for always believing in me. Thank you to my advisor Dr. Bauer and committee members, Dr. Rogers and Dr. Oxley, which gave me the benefit of their priceless expertise. I need to thank our department's research associate, Trevor Bihl, whose partnership brought imaginative ideas and solutions. I would also like to thank the Air Force Research Laboratory for generously sponsoring this research. I would also like to thank Dr. Takeo Kanade and Dr. Louis Denes of Carnegie Mellon University for generously sharing their hard work and the CMU Database. To everyone mentioned, I thank you for the most exciting and rewarding educational experience of my life.

Table of Contents

Page

Abstract iv

Acknowledgements iv

Table of Contents v

List of Figures 1

List of Tables 8

List of Equations 9

I. Introduction 11

II. Human Recognition 14

Cognitive Research 14
Cognitive Disorders 22
Human vs. Computer Recognition 24
Summary 25

III. Face Recognition Systems 27

Research Trade Space 28
Common Algorithms 30
Eigenface 30
Fisherface 36
Algorithm Extension for Pose and Internal Features 38
Neural Networks 41
Face Space 44
Distance Metrics 47
Biometric Features 51

Soft Biometrics 51
Spectral Properties 55
Recognition Challenges 63
Occlusion 63
Illumination 66
Pose 70
Circumvention 76
Hyperspectral Face Recognition 82
Hyperspectral Research 83
Fusion 88
Fusion Strategies 89
Theoretical Framework 94
Value of Diverse Classifiers 96

IV. Research 99

Image Data 100
QUEST Methodology 104
Illustrative Example 107
Research Goals 113
Data Exploration 113
Preprocessing 116
Spatial Recognition 123
Spectral Recogntion 126
Interest Point Recognition 135
Published Algorithm Comparison 139
Algorithm Enhancements 141
Adaptive Feedback 172

V. Results and Findings 181

VI. Conclusion 197

VII. Future Work 203

Bibliography 210

List of Figures

Figure 1: Blurred Images of Familiar Faces [4] 14

Figure 2: Compressed Images of Familiar Faces [4] 15

Figure 3: Negative Contrast of the Album Cover, "We Are the World" [4] 15

Figure 4: High Frequency Spatial Information [4] 16

Figure 5: Holistic Processing of Faces [4] 17

Figure 6: Combinations of Internal and External Features [6] 17

Figure 7: Recognition Results of Internal, External and Whole Face [6] 18

Figure 8: Observed Face, Average Face, and Caricatured Face [4] 19

Figure 9: Face Space Depicting Aftereffects [4] 19

Figure 10: Face Distribution of Faces [8] 21

Figure 11: ROC of Human and Computer Performance on Matching Faces [15] 25

Figure 12: Conventional Biometric System Components 27

Figure 13: Face Recognition Research Trade Space 30

Figure 14: Basis of Eigenfaces 33

Figure 15: Summary of Eigenface Procedure [20] 36

Figure 16: Same Person with Different Expression and Lighting [22] 36

Figure 17: Eigeneyes, Eigennose, and Eigenmouth [24] 39

Figure 18: Eigenfeature Testing [24] 40

Figure 19: Recognition Rate for Eigenfaces, Eigenfeatures and Combined [24] 40

Figure 20: Diagram of RBF Neural Classifier [27] 42

Figure 21: Recognition Capability with Increasing Number of Eigenfaces [21] 46

Figure 22: Optical Properties of the Skin [48] 55
Figure 23: Variations in Spectral Remittance of Skin [48] 56
Figure 24: Differences in Mean Spectral Reflectances [47] 57
Figure 25: Predicted Spectral Reflectance based on UV Exposure [49] 58
Figure 26: Spectral Signature of Lips (Green Line) Using Visible Wavelengths 59
Figure 27: NIR Spectral Signature of Hair vs. Skin 61
Figure 28: Spectral Signatures of Various Hair Colors [54] 61
Figure 29: Face Representations without Eyes and Eyebrows [4] 62
Figure 30: Face Images (Undistorted, horizontal, and vertical occlusions) [21] 64
Figure 31: Recognition Rates with Increasing Occlusion Size [21] 64
Figure 32: Original and Blurred Images [21] 65
Figure 33: Recognition Rates for Blurred Images [21] 65
Figure 34: Adjusting Illumination with Discrete Cosine Transform [57] 67
Figure 35: Global Irradiance Spectra Measured at Different Times [58] 68
Figure 36: Spectral Signature of Forehead under Different Illumination [58] 69
Figure 37: Producing the Set of Scaled Images [61] 72
Figure 38: Minima and Maxima of the DoG Images [61] 72
Figure 39: Creating the Keypoint Descriptor [61] 73
Figure 40: Scale Invariant Feature Transform (SIFT) - Key Point Detection 74
Figure 41: SIFT Applied to Upright and Inverted Images 75
Figure 42: SIFT Applied to Scaled and Rotated Image 75
Figure 43: National Geographic Face Images for Circumvention [65][66] 76

Figure 44: NIR Skin Detection Compared to Doll's Face 77
Figure 45: NIR Skin Detection When Make-up is Applied 79
Figure 46: Vein Patterns in Hyperspectral Hand Image 80
Figure 47: Vein Patterns on Hand and Face 80
Figure 48: Spectral Signature of Various Individuals Faces and Regions [93] 84
Figure 49: Spectral Angle Comparison of Eight Test Subjects [93] 85
Figure 50: Data Fusion Using IR and MMW Images [104] 90
Figure 51: Borda Count Method 94
Figure 52: UCI Database – Face Image with Expressions and Rotation [96] 100
Figure 53: CMU Hyperspectral Database 101
Figure 54: Equinox Multispectral Face Database [99] 102
Figure 55: Oulu Color Image Database [100] 103
Figure 56: Agents and Level of Abstraction [102] 105
Figure 57: Skin Component for Face Detection (Theoretical) [106] 108
Figure 58: Hyperspectral Data Cube Covariance Matrix 108
Figure 59: First Principal Component of HSI 109
Figure 60: Additional Principal Components and Spectral Signatures 110
Figure 61: Initial Hyperspectral Face Recognition Process 111
Figure 62: Blue and Green Wavelengths of an Image 114
Figure 63: Variability of 17x17 Pixel Spectral Sample from Forehead 115
Figure 64: Movement of Subject during Image Capture 116
Figure 65: Appearance Changes between Training and Test Sets 116

Figure 66: Preprocessed Image Used in the CSU Face Evaluation System [36] 118
Figure 67: Grayscale Images before Preprocessing..... 121
Figure 68: Training Subjects..... 121
Figure 69: Grayscale Images after Preprocessing..... 122
Figure 70: Training Images after Preprocessing..... 122
Figure 71: Initial Eigenface Testing Before Preprocessing 123
Figure 72: Eigenface Testing - Retaining Varying Number of Eigenfaces 124
Figure 73: Eigenface Testing - Single Wavelength vs. Average Image 125
Figure 74: CMS Using Eigenface on Manually Preprocessed Faces 126
Figure 75: Manual Selection of Points for Spectral Matching 126
Figure 76: Spectral Angle Matching from Selected Tissue types..... 127
Figure 77: Sum Fusion of Spectral Signatures 128
Figure 78: Clusters Shown Separately Using K-means Algorithm [97]..... 129
Figure 79: K-means Clustering Results 130
Figure 80: Additional K-Means Testing and Different Illumination..... 131
Figure 81: Elbakary's Published Results [97] 132
Figure 82: Pan's Published Results Using Spectral Signatures [58]..... 133
Figure 83: K-means and Spectral Matching Testing 134
Figure 84: Application of SIFT to Determine Interest Points and Matches 136
Figure 85: SIFT Experimental Results with Varying Distance Thresholds 137
Figure 86: SIFT Recognition Errors 138
Figure 87: Spectral Face Implementation for CAL and CMU Databases 139

Figure 88: Testing Results of Select Algorithms on CMU Data 140
Figure 89: Contextual Layers of Skin and Hair 144
Figure 90: Using Skin Detection to Determine Centroid 145
Figure 91: Edge Detection Efforts 146
Figure 92: Incorporation of Edge Information for Recognition 147
Figure 93: First Five Principal Components: CMU Data (Top) and CAL Data (Bottom) 148
Figure 94: Application of BACON Outlier Detection for Hair Signatures 154
Figure 95: Performance Impact of Applying Outlier Identification 155
Figure 96: Improvement in Biometric System Using Soft Biometrics [38] 156
Figure 97: Spectral Angle Distances 157
Figure 98: Joint Directors of Laboratories Data Fusion Model [104] 159
Figure 99: Hyperspectral Face Recognition Hierarchy 161
Figure 100: Mathematical Formulation of HSI FR Problem 163
Figure 101: Applied HFR Hierarchy 166
Figure 102: Score and Rank Fusion Results for HFR 167
Figure 103: CMS Plot of HFR Methodology vs. Competing Methods Tested on CMU Data 168
Figure 104: CMU Validation Data Set 169
Figure 105: CMU Validation Results 169
Figure 106: Recognition Agent Interface 171
Figure 107: Adaptive Logic 174

Figure 108: Extension of Biometric Research 175
Figure 109: Ensemble Enhancement for Future Research 178
Figure 110: HSI FR GUI for Data Exploration and Strategy Development 179
Figure 111: Cumulative Match Score Results Including Score Fusion 181
Figure 112: Box Plot of Match Scores for All Probes 183
Figure 113: Box Plot of Match Scores for Single Probe 184
Figure 114: Visual Depiction of Match Candidates 185
Figure 115: Distribution of Match Scores (Global and Local) 186
Figure 116: GUI Evaluation of Sub-standard Matches (Subject 12) 187
Figure 117: Distribution of Substandard Match (Subject 12) 188
Figure 118: GUI Display after Reduction Step - Negligible Improvement for Subject 12 188
Figure 119: Box Plot and Distribution of Matching Scores - For First Reduction for Subject 12 189
Figure 120: Box Plot for Evaluation of Reduction Set 189
Figure 121: GUI Display for Second Reduction - Significant Improvement for Subject 12 190
Figure 122: Box Plot and Distribution for Second Reduction (Subject 12) 190
Figure 123: Box Plot for Final Reduction Evaluations 191
Figure 124: CMS for Various Fusion Combinations and Weightings 192
Figure 125: Global Box Plot of All Match Scores for 6 Agent Unity Weighting 193

Figure 126: Progression of Score Distributions for Subject 36 (6 Agent/Unity Weighting) 194
Figure 127: Global Box Plot of All Match Scores for 7 Agent SIFT Biased Weighting..... 195
Figure 128: Progression of Score Distributions for Subject 36 (7 Agent/SIFT Biased Weighting) 195
Figure 129: Recognition Progression for Subject 30 (7 Agent SIFT Biased) 196
Figure 130: QUEST Situational Awareness [120]..... 198
Figure 131: Disguise Raises Security Concerns [125] 201

List of Tables

Table 1: NDSI Values for Various Materials [68] 78

Table 2: Ear Recognition Studies [70] 81

Table 3: Summary of Hyperspectral Face Recognition Research Efforts 87

Table 4: Contents of CMU Database 120

List of Equations

Equation 1: Matrix of Gallery Faces - Difference from Average Face 31

Equation 2: Determination of PCA Eigenvectors and Eigenvalues [20] 32

Equation 3: Alternative Representation of Covariance Matrix of Faces 32

Equation 4: Relationship between Eigenvectors and Eigenvalues 33

Equation 5: Eigenfaces 33

Equation 6: Projection into Face Space 34

Equation 7: Distance between Faces 34

Equation 8: Face Space Threshold 35

Equation 9: Between and Within Class Scatter [22] 37

Equation 10: Definition of Distance Metric 47

Equation 11: Common Distance Metrics 48

Equation 12: Correlation Distance Measurement [36] 48

Equation 13: Covariance Distance Measurements [36] 49

Equation 14: Relation between PCA/LDA and Mahalanobis Space 49

Equation 15: Mahalanobis Distance Measurements [36] 50

Equation 16: Additional Distance Measurements [37] 51

Equation 17: Scale Function [60] 70

Equation 18: Variable Scale Gaussian Function [60] 71

Equation 19: Difference of Gaussian Filter [62] 71
Equation 20: Difference of Two Nearby Scales [61] 71
Equation 21: Gradient Magnitude and Orientation Calculation [61] 73
Equation 22: NDSI and NIMI Calculation [68] 78
Equation 23: Spectral Angle [93] 84
Equation 24: Mahalanobis Cosine Distance [36][75] 91
Equation 25: Sum of Mahalanobis Cosine Distances [75] 91
Equation 26: Probability of Fusing Biometric Features [80] 95
Equation 27: Probability of *k* Errors for Majority Fusion Rule 95
Equation 28: Likelihood of the Majority Rule Error 96
Equation 29: Error Cost for the Optimum Bayesian Fusion of N Sensors [84][85] 96
Equation 30: NDSI Substitute Approach 142
Equation 31: NDVI Calculation 143
Equation 32: Mahalanobis Distance [113] 152
Equation 33: Mahalanobis Distance of Basic Subset [113] 153
Equation 34: Formulation of Correction Factor [113] 153

I. Introduction

Social interaction and communication depends heavily on the amazing face recognition capability that humans possess. In a myriad of environments and views, people are able to quickly recognize and interpret visual cues from another person's face. This impressive ability has sparked the interest of researchers from the cognitive sciences to statistical pattern recognition. This remarkable aptitude is the elusive performance standard that motivates and confounds developers for computer vision and biometric recognition systems.

During the last decade, there have been several factors that have accelerated the advances in face recognition and biometric technologies. One factor is the intensified focus on security issues throughout the world because of the expanding threat and terrible repercussions of terrorist acts. A second factor is the advancement and availability of supporting technologies. The popularity of portable electronic devices such as the cell phone, personal computing devices, and digital cameras create the means for a very capable surveillance system for any office or street corner. This technology along with the growing wireless network can enable the persistent monitoring of a very large portion of the globe.

Within this environment, face recognition offers an attractive and non-intrusive biometric that can be leveraged to exploit such opportunities. Unfortunately, the task of

developing a robust recognition system is not a trivial problem. These systems must possess the sensitivity to detect the smallest changes in human appearance and a robustness to be used in a multitude of operating environments but still maintain an efficiency that can be utilized real-time on large segments of the world's population. Despite years of creative development and scientific advancement, current face recognition algorithms are challenged by natural variations in illumination, pose, and expression and will soon face an expected barrage of spoofing attempts in critical security applications.

Given these challenges, the design of a biometric identification system needs to possess certain characteristics to make it an effective operational system. These attributes include universality, distinctiveness, permanence, collectability, performance, acceptability, and circumvention [1]. The face recognition modality unfortunately suffers from weaknesses in the areas of uniqueness, performance, and circumvention [2]. The well-known biometric researcher and educator, Jain [3] lays out three basic requirements for a face recognition system to be effective. These basic capabilities are an ability to detect whether a face is present, the means to locate the face, and then the capacity to recognize the face from a general viewpoint. Within this structure, the performance target for a computer based face recognition application is to mitigate these weaknesses while achieving a recognition capability equal to that of a human.

The use of hyperspectral imagery (HSI) and the contextual information layers contained within these image cubes provides the cues to creating a hierarchal methodology that can address the common challenges for face recognition systems. Using a variety of features that play an important role in human cognition and span

general characteristics to specific attributes, a fusion hierarchy is developed that incorporates key QUEST (QUalia Exploitation of Sensor Technology) tenets. The result is a face recognition methodology that provides a performance improvement and operational robustness that exceeds classical face recognition methodologies.

This paper documents the research to address the weaknesses of classic face recognition algorithms with the development of a novel methodology and the utilization of hyperspectral imagery. In Chapter 2, the insights from the study of human recognition is examined to discover clues that may valuable to the development of a similar capability using computer recognition algorithms and methods. Chapter 3 follows with a discussion of the more common face recognition algorithms. Important design aspects are reviewed to include the construction of the comparison space, selection of distance measures and possible strategies for improving the performance of existing recognition algorithms. Chapter 4 catalogues the research process and challenges faced during the exploration of data and development of the QUEST face recognition methodology. The experimental results and performance comparisons are contained in Chapter 5. The conclusion in Chapter 6 summarizes many of the insights gleaned from this effort as well as the contributions of this research. The promise of future research in this field is immense and Chapter 7 discusses some of the possibilities and associated security and defense applications.

II. Human Recognition

Cognitive Research

A compilation of lessons learned from cognitive science was published by Sinha [4] and provides a valuable foundation for researchers in the field of computer vision. The standard for face recognition performance is to match or exceed the ability of humans to recognize faces and so leads us to this natural and logical starting point. In the design of current face recognition systems, there is usually an effort to process imagery with ever-increasing resolution in order to improve performance. As illustrated in Figure 1, only a small amount of resolution is required to recognize well-known faces. In this instance, the blurred images, equivalent to an image resolution of only 7x10 pixels, are recognizable to most readers despite the relatively poor picture quality [4].

Figure 1: Blurred Images of Familiar Faces [4]

A person's robust capability to identify an object is not solely reliant on the quality of the image. Another example of degradation that human recognition can overcome is illustrated in Figure 2 [4]. In these images, the width of the face is

compressed to one quarter of the original size and yet these familiar faces are easily recognized. In this case, much of this ability is attributed to the fact that the proportions are stable in at least one direction (vertically).

Figure 2: Compressed Images of Familiar Faces [4]

The amount of variation that a human recognition system can withstand is impressive, but there are some important visual cues that can quickly result in a severly degraded system. By altering the color and pigmentation from the album cover of "We Are the World" (Figure 3) many famous and recognizeable performers such as Micheal Jackson, Stevie Wonder, and Ray Charles become almost impossible to identify. The change in contrast, similar to that seen in photograph negatives, offers a confusing representation from which many observers have trouble identifying a single face from the many celebrities located in the crowd [4].

Figure 3: Negative Contrast of the Album Cover, "We Are the World" [4]

Sinha [4] offers another illustration to convey the importance of contextual information contained in color or grayscale images. In of many recognition algorithms, considerable effort is applied to obtaining detailed edge information. In Figure 4, a contour or edge map representation provides a confusing representation of two well known actors, Jim Carrey and Kevin Costner [4]. Advancements in the detection and the retention of high frequency information has benefited face recognition systems but this representation alone appears to be insufficient for our own human recognition system.

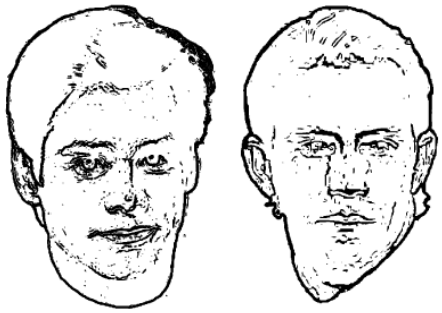

Figure 4: High Frequency Spatial Information [4]

An important element from cognitive research that has influenced the development of face recogniton algorithms is the role that holistic processing plays in our own recogntion system. As depicted in Figure 5, the image halves of Woody Allen and Oprah Winfrey are not easily discernible until they are seperated into distinct halves. The natural holistic processing of the whole image is difficult to overcome and impedes the ability to distinguish the two identities [4]. By trade, portrait artists are intimately familiar with the features of the face and are suspectible to the holistic influence of the face. In the instruction of new artists and in the practice of accomplished artists, a potrait

begins by first schetching the silhouette of the head followed next by the placement and construction of internal features all guided by the context of the face's outline [5].

Figure 5: Holistic Processing of Faces [4]

Jarudi's research explored the relative importance of these external features, such as the outline or shape of the head and face, versus the internal features of a face, to include the eyes, nose and mouth, as well as their configuration [6]. Using human test subjects to evaluate each feature hierarchy, it was discovered that the importance of these features vary as the resolution of the image changes. An example of the images used to test these effects are shown in Figure 6.

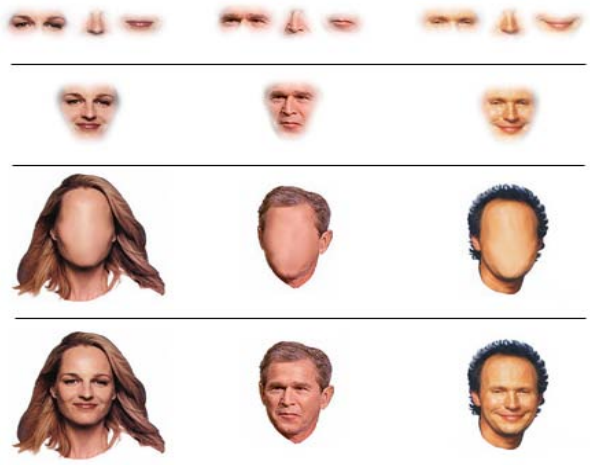

Figure 6: Combinations of Internal and External Features [6]

The results, shown in Figure 7, indicate that the recogniton of external features decrease at a gradual linear rate as the resolution decreases, whereas the recognition rates for internal features decreases non-linearly and more abruptly as resolution decreases. These findings suggest that image quality or resolution should guide the relative weightings for external features versus internal features for classification systems that integrate both for maximum peformance.

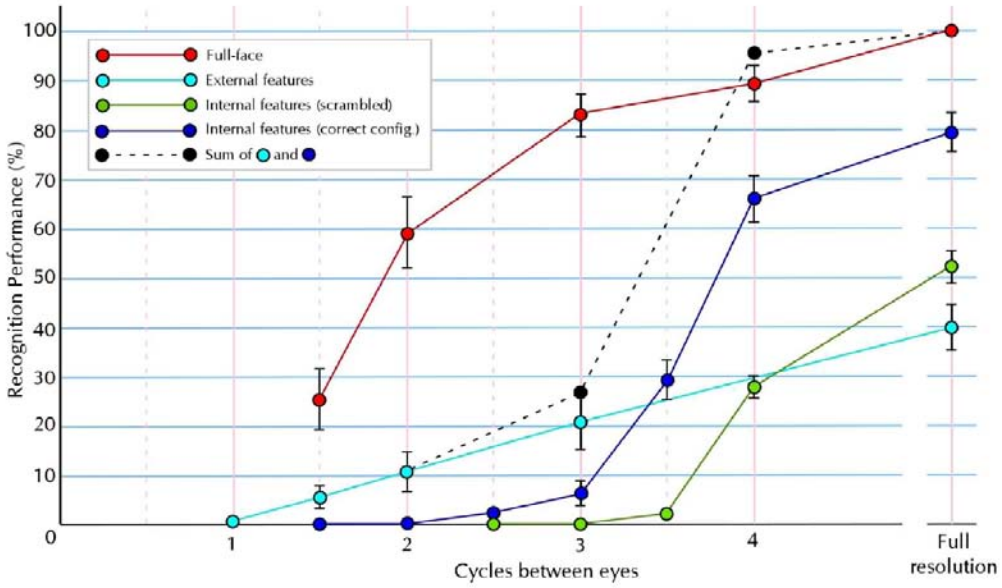

Figure 7: Recognition Results of Internal, External and Whole Face [6]

In addition to these insights concerning the processing of images and their features, current cognitve research can help strengthen our understanding of how faces are stored and recalled from memory. Studies reveal that a caricatured representation of the true image can be beneficial for human recognition performance [4]. A veridical or true representation of a face is important in correctly matching the identity of an

individual but an exaggerated version, emphasizing changes from an average or generic face, has been shown to support recognition accuracy. An example of this representation with respect to face shape and pigmentation is shown in Figure 8. This figure depicts the true representation, the average face of all those in memory, and a caricatured representation embellishing the difference or variance between the two [4].

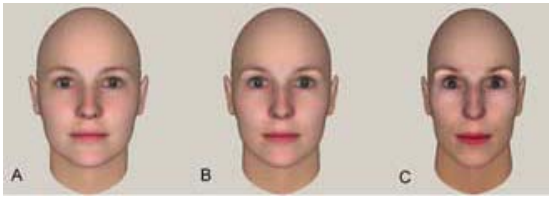

Figure 8: Observed Face, Average Face, and Caricatured Face [4]

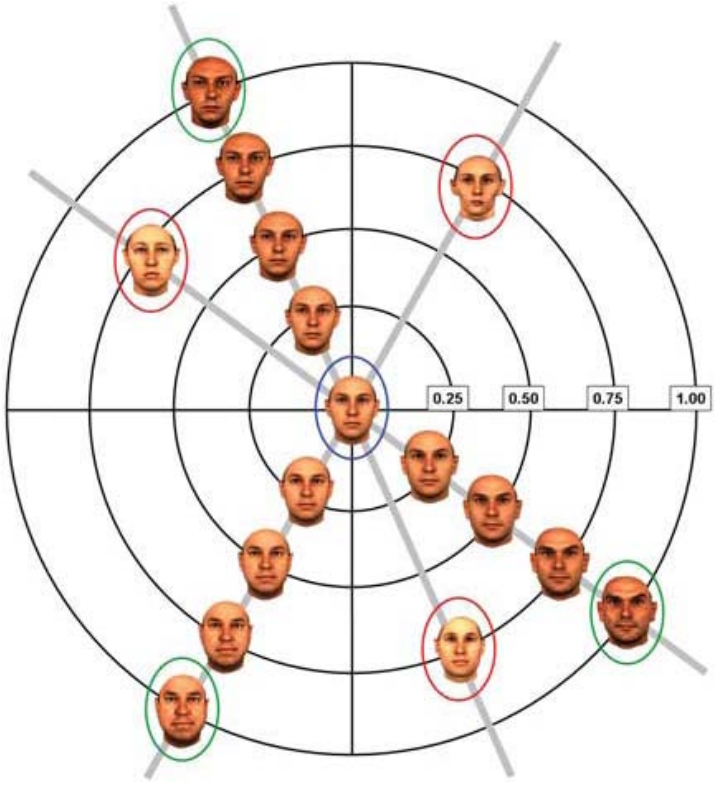

Figure 9: Face Space Depicting Aftereffects [4]

In addition to the impact that variance has on face representations, there is also a detectable bias that is obtained after prolonged observation of a particular face [4]. This concept is illustrated in Figure 9, as the center image circled in blue is the true image of the person we initially observed, and later want to identify. After shifting focus to a face located on the perimeter, annotated by a green circle, and maintaining the gaze for a sustained period of time, our recollection of the original face is now biased in the opposite direction, away from the face image on the perimeter and along the radial axis to a representation marked by a red circle.

In one study, the recognition capability of humans was investigated to better understand the distribution of faces that are stored in memory as well as the dimensionality of the human face space model [7], [8]. The findings of this research, not surprisingly, showed that distinctive faces are more easily identified than typical faces and intuitively this may be due to the location and distribution of these faces in the human face space. During testing, distinctive faces resulted in both a higher true positive rate and a lower false positive rate than typical faces. This discovery parallels a phenomenon known as the Doddington Zoo effect identified in voice identification [9]. In much the same way, easily distinguishable individuals, known as sheep, had a higher ability to be accurately identified, while other more typical voices, known as goats, were consistently difficult to recognize.

Using a comparison space or face space perspective, these distinctive faces are located further away from neighboring faces and are more easily identified. The center of the space represents the average face of the population with typical faces located densely near the center and a sparse distribution of distinctive faces located on the

peripheral. An assumption of a multivariate normal distribution of faces is commonly used to represent this distribution of face images [7]. This distribution was demonstrated empirically by Johnston's research [8] and the results are portrayed in two dimensions in Figure 10. The similarity between faces is measured by the separation distance and the summed similarity between all exemplars can serve as an indication of the expected identification performance. This insight seems to be understood by the academic research community, as many common face recognition databases are comprised of a varied and diverse group of subjects. If the current limitations to understanding the dimensionality of the human face space were surpassed, algorithms and metrics could be better tailored to mimic the ability to measure and distinguish features in this multi-dimensional representation space.

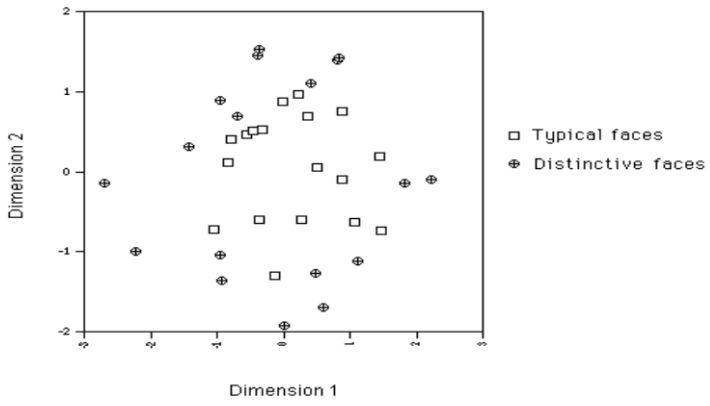

Figure 10: Face Distribution of Faces [8]

Cognitive Disorders

Valuable insights are gained by exploring the normal performance and representations used in human recognition, but also in the dual of the problem by examining disorders that can destroy this fragile ability. Two cognitive conditions that affect recognition are Prosopagnosia [11] and the Capgras delusion [10]. The breakdown in the critical links affected by these disorders may offer clues to the structure and functionality of the systems that provide robust recognition capability.

Prosopagnosia is the failure of people to recognize faces very familiar to them. People with this disorder maintain the ability to interpret facial gestures but a close friend can become an instantaneous stranger with as little as a change in lighting or background setting [11]. On the other hand, Capgras delusion is the belief that a face, although easily recognized, is a disguised imposter or body double [10]. These two disorders and their underlying symptoms, however different, may collectively provide some understanding of the cognitive process used during human recognition.

For years, human recognition was believed to be a sequential process. However, the close examination of the skin conductance response (SCR) of patients afflicted with these disorders suggests otherwise [10]. SCR is a measure of the skin's electrical conductance caused by changes in the level of perspiration in the sweat glands due to nervous system activation [10]. Prosopagnosia patients show indications of elevated SCR when they look at a familiar face, indicating an underlying positive response by the nervous system, despite the lack of a conscious recognition of the person. Alternatively, a patient suffering from the Capgras delusion shows no changes in their SCR, despite the

fact that can outwardly name the face they are looking at. In both these cases, the ability to recognize a face is uncoordinated at the emotional or physiological level. A fully functional system relies on the integration and synchronization of both these systems.

Tranel [11] found that patients incapable of face recognition were still able to recognize facial gesture, age, and gender. This outcome once again suggests that there are separate processes involved, an overt process that enables humans to identify faces and a covert process that allows individuals to connect familiar information for confirmation [10], [11]. These separate processes may also mirror the cognitive and Libet processes discussed later when the Qualia Exploitation of Sensor Technology (QUEST) methodology is discussed. The inability to integrate these processes results in significant capability degradation for the human recognition system and perhaps the failure to consider these links for computer recognition systems prevents them from reaching the performance benchmarks seen in human recognition.

Environmental variations can also have a significant effect on the human recognition system. Yin [12] explored the degraded ability of humans to recognize faces when they are upside down. Findings from his research demonstrated that many mono-oriented objects are difficult to recognize when inverted and that the human face is one that is most predominately affected. The well-documented Thatcher illusion [13] shows the dramatic effect that inverting a face has on our powers of observation. In the Thatcher illusion, the observer is unable to realize that internal facial features, such as the mouth and eyes, have been flipped upside down when portrayed in a face that has itself been inverted.

This powerful effect of face orientation has been seen firsthand in the U.S. Space Program where astronauts merely observe another crewmember's upside down face and the entire perception of their orientation is quickly altered often resulting in disorientation and sometimes motion sickness [14]. This particular shortcoming of the human recognition system can be overcome by using pose invariant computer algorithms such as Scale Invariant Feature Transform (SIFT) that can identify objects regardless of orientation.

Human vs. Computer Recognition

The ability of the human recognition system is held as the performance benchmark for classification methods and was challenged for the first time in the 2006 Face Recognition Vendor Test (FRVT). The FRVT tested human performance against the best performing commercial algorithms [15]. The human recognition evaluation used 26 students who had to evaluate 80 pairs of male and female faces. The image pairs were presented side by side for only two seconds. Afterwards, students selected a rating from 1 (sure both images were the same person) to 5 (sure both images were different people). The results, shown in Figure 11, showed that several of the more advanced systems could exceed human ability in this subjective test [15]. There was no mention whether computational algorithms used the same incremental evaluation scale.

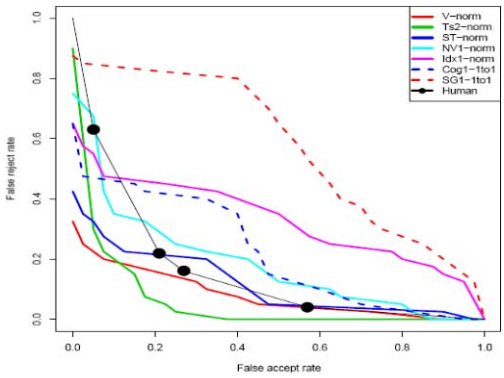

Figure 8: ROC of human and computer performance on matching faces across illumination changes. ROCs for algorithms in Figure 7 are plotted. The ROC plots FAR against FRR. Perfect performance would be the lower left hand corner (FAR=FRR=0).

Figure 11: ROC of Human and Computer Performance on Matching Faces [15]

Although commercial recognition systems appear to be approaching the ability of human recognition, these test results were obtained in a controlled environment. In addition, the subjective scale and time constraint used in testing provides an advantage to the recognition algorithms. Unfortunately, these same recognition algorithms are required to operate in uncontrolled environments challenged by various levels of cooperation from the intended subjects. Despite the level of sophistication of these commercial systems, the performance and reliability of these systems still require a human in the loop to confirm the results for real world applications.

Summary

Cognitive research has greatly aided our understanding of the human recognition system and can serve as a useful guide for recognition system development. Using these findings, the training procedures and transformation techniques used to create a face

space could exploit the variance and bias of observed faces as well as using the relative value of shape, color, and resolution. Face recognition algorithms may benefit from a structure that represents a parallel neural system of the human mind that incorporates holistic face recognition processes, feature based representation, and various levels of semantic representation. Research has shown that recognition processes that fail to combine both spatial and contextual information have limited capability for the demanding task of face recognition.

III. Face Recognition Systems

In an attempt to introduce several important considerations for the design and implementation of a face recognition system, an assortment of topics will be covered in this chapter. This discussion will start with the basic components of a generic biometric system and move to the research trade space in which this effort will reside. A sampling of common algorithms will be reviewed for consideration to include Eigenfaces, Fisherfaces, and neural networks. Additional considerations within these algorithms, such as the comparison space, distance measurements and feature selection, play an important role and need some discussion as well. With this basic understanding, a summary of current challenges for recognition systems will be examined.

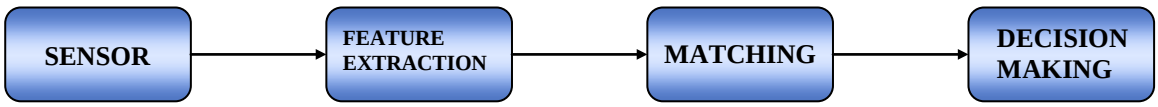

Figure 12: Conventional Biometric System Components

Biometric systems are comprised of four basic components. The components, depicted in Figure 12, are the sensor, feature extraction module, matching module and the decision-making module [16]. The sensor module acquires the biometric data (i.e. a measurement) from the intended subject. The feature extraction module processes the captured data from the sensor and extracts features. The matching module compares the extracted features against stored features that are saved in memory and generates

comparisons called match scores. Match scores are comparisons made in a multi-dimensional comparison space and are a measure of distance between two images. The decision-making module takes these scores and determines the user's identity by selecting the stored features (identification) associated with the smallest match score or by evaluating the obtained match score against a threshold for the claimed identity's features (verification).

Research Trade Space

This research will utilize existing hyperspectral data that has been obtained with an experimental sensor. This research effort will focus on the last three components, particularly the selection of features and how they are utilized. For face recognition systems, feature extraction can be based on internal features, a holistic representation, or a combination of both. Internal features range from the shape and size of the eyes, nose, or mouth to their location on the face as well as their geometric proximity to each other. A holistic representation captures the entire image using the individual parts and the surrounding backdrop of the entire face. This research will explore a range of these spatial features and look for opportunities to combine them for synergistic effects.

The spatial domain has been the predominant dimension of face recognition research with a primary focus on grayscale and color images. More recently, the advancement of sensor technology has allowed examination of a wider continuum of the electromagnetic spectrum. Thermal imagery has proved beneficial for detection and tracking functions but for identification purposes, it suffers from sensitivities to changes

in environmental temperature, thermal patterns of the face, and its inability to penetrate glass [17], [18]. Due to national security requirements, wavelengths beyond the visible and infrared (IR) are being explored for use in security systems to penetrate clothing and skin. A recent and highly controversial example of this technology has been seen in the terahertz and millimeter wave technology being tested and implemented in many of our major airports.

Finally, the widespread use of video imagery in security systems has resulted in the need for reliable and efficient methods to process the large amount of collected data. The performance requirements for these methods must exceed the processing capability of human beings given the volume of data and number sensors. These processing systems need to incorporate methods developed to handle both the spatial and spectral elements of the data, offering an efficient way ahead to take advantage of the temporal aspects of real time multispectral video. This same challenge of processing vast amounts of data is facing our nation's military forces as full motion video and advanced sensors are employed to meet the targeting requirements of current operations.

The method of analyzing extracted features and images is dependent on the chosen algorithm, the projected comparison space, and the metrics selected for assessment. This daunting assignment and the aim of this research is to take this general task and apply it to the spatial, spectral, and temporal domains simultaneously. The development of an approach that integrates all three dimensions can mitigate many of the challenges and offer new capabilities not seen in any single domain. The goal of this study is to provide a framework to process and match imagery using cues from the research environment shown in Figure 13.

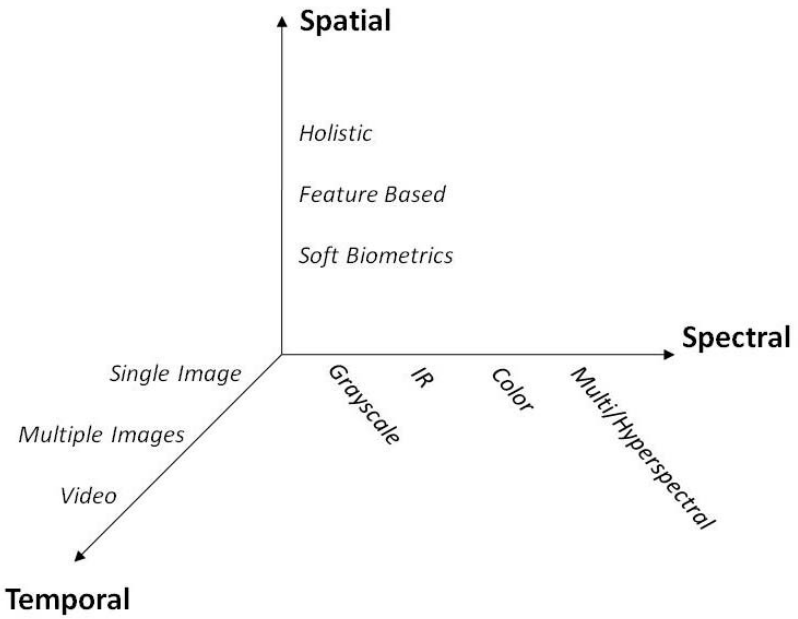

Figure 13: Face Recognition Research Trade Space

Common Algorithms

Before creating a methodology for hyperspectral face recognition, it is important to understand the theory behind some of the most accepted algorithms used in the spatial and spectral dimensions. By identifying and combining the strengths of individual methods for each dimension, a more capable methodology should be possible. A brief overview of two of the leading face recognition algorithms follows.

Eigenface

One of the most popular and widely tested face recognition techniques is the Eigenface method. Kirby and Sirovich [19] introduced the concept of representing faces

with a weighted sum of eigenvectors almost two decades ago. Shortly after, Turk and Pentland [20] used the concept to devise the Eigenface method. This holistic approach was as an attempt to replicate the cognitive aspects of human recognition. At the time, this method provided an alternative to the many feature-based methods that relied on specific features of the face, but discarded a large portion of the image as well as the contextual information contained therein.

A brief description of the eigenface algorithm follows [20], [21]. For each of the *M* training face images, comprised of *N* x *N* pixels, the pixel intensity values from each image (*i*) are concatenated into a single vector, Γ_i. The average face image, Ψ, is calculated using all *M* face vectors. The difference from the average face, Φ_i, is then calculated for each image and then stored into a matrix of difference face vectors, $\boldsymbol{A}$. These steps are summarized in Equation 1.

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i$$

$$\Phi_i = \Gamma_i - \Psi$$

$$\boldsymbol{A} = [\Phi_1, \Phi_2 \ldots \Phi_M]$$

Equation 1: Matrix of Gallery Faces - Difference from Average Face

When principal component analysis (PCA) is applied, it is used to find a set of *M* orthonormal vectors, *u*, that best represents the data distribution. These M vectors are chosen to satisfy the following equation.

$$\lambda_k = \frac{1}{M}\sum_{n=1}^{M}(u_k{}^T \Phi_n)^2$$

Equation 2: Determination of PCA Eigenvectors and Eigenvalues [20]

These scalars, λ_k, and vectors, u_k, are the eigenvalues and eigenvectors of the covariance matrix (***C***). Unfortunately, this becomes computationally prohibitive due to the size of the images (N x N) and the resulting size of the covariance matrix (N^2 x N^2). A more useable representation, ***L***, utilizes the same eigenvectors but in a compact structure (M x M). The calculation of ***L*** using the matrix of difference face vectors, ***A***, is shown in Equation 3.

$$\boldsymbol{C} = \tfrac{1}{M}\sum_{i=1}^{M}\boldsymbol{\Phi}_i\boldsymbol{\Phi}_i^T = \boldsymbol{A}\boldsymbol{A}^T$$

$$\boldsymbol{L} = \boldsymbol{A}^T\boldsymbol{A}$$

Equation 3: Alternative Representation of Covariance Matrix of Faces

Using the eigenvalue (μ_i) and eigenvector $(\boldsymbol{v}_i)$ of ***L***, the eigenvalue (μ_i) and eigenvector $(\boldsymbol{A}\boldsymbol{v}_i)$ comprised of $\boldsymbol{v}_i$, can be derived for ***C***. The l eigenvectors of ***L*** are used to represent the distribution of the data but in a smaller dimensional subspace, known as the face space. The relationship between the eigenvalues and eigenvectors of ***C*** and ***L*** are shown in Equation 4.

$$\boldsymbol{L}\boldsymbol{v}_i = \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v}_i = \mu_i\boldsymbol{v}$$

$$\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v}_i = \boldsymbol{A}\mu_i\boldsymbol{v}_i$$

$$\boldsymbol{C}\boldsymbol{A}\boldsymbol{v}_i = \mu_i\boldsymbol{A}\boldsymbol{v}_i$$

Equation 4: Relationship between Eigenvectors and Eigenvalues

The eigenvectors (v_i) of $\boldsymbol{L}$ are used to produce the basis of ghostly face images $(\boldsymbol{u}_l)$ that are illustrated in Figure 14.

Figure 14: Basis of Eigenfaces

These images, known as eigenfaces, serve as the basis for our face space and are calculated in Equation 5.

$$\boldsymbol{u}_l = \sum_{k=1}^{M} \upsilon_{lk}\boldsymbol{\Phi}_k$$

Equation 5: Eigenfaces

A linear combination of these eigenfaces is used to recreate all the faces stored in the gallery. A new test face, called a probe, is similarly transformed to an average subtracted face ($\boldsymbol{\Phi}_i$) much like the gallery images were previously. The probe $(\boldsymbol{\Gamma}_{new})$ is

projected into a face space and represented by a vector of weights ($\mathbf{\Omega}^T$) that describe the contribution from each of the eigenfaces. Equation 6 formulates these steps.

$$\omega_i = \mathbf{u}_i(\Gamma_{new} - \Psi)$$

$$\mathbf{\Omega}^T = [\omega_1 \omega_2 ... \omega_M]$$

Equation 6: Projection into Face Space

To further reduce the computational requirement, a subset *k* of the initial *M* eigenfaces are usually retained for use. These *k* eigenfaces are chosen to represent a majority of the information in the face space based on their eigenvalues. This is no different than deciding on how many significant principal components to retain during PCA to capture the majority of variability in any generic dataset. Matching the identity of the new test face is accomplished by finding the minimum distance (ε_k), in this case Euclidean distance, between the probe (Ω_{new}) and a gallery face (Ω_k) that have both been projected into the comparison face space (Equation 7).

$$\varepsilon_k = \sqrt{\left\| \mathbf{\Omega}_{new} - \mathbf{\Omega}_k \right\|^2}$$

Equation 7: Distance between Faces

This approach can also be applied to discern face images and from non-face images, or the recognition of any object given an adequate gallery of like object images. When a face representation uses a subset of *k* eigenfaces, some information is lost and this quantity is referred to as reconstruction error. The radius of the initial face space can serve as a threshold to assist in comparing the magnitude of these errors. Based on this

initial face space threshold, an image's reconstruction error would indicate that it is a non-face image if its error is larger than the face space threshold. This application can be applied in the following manner to filter out any images that are not faces, avoiding unnecessary processing, and matching attempts of non-face images (Equation 8).

$$\Phi_{reconstruction} = \sum_{i=1}^{M} \omega_i u_i$$

$$\theta_{face\ threshold} = \frac{1}{2}\max(\sqrt{\left\|\Omega - \Omega_k\right\|^2})$$

$$\varepsilon^2 = \left\|\Phi_{image} - \Phi_{reconstruction}\right\|^2$$

$$\varepsilon > \theta_{face\ threshold}$$

Equation 8: Face Space Threshold

Likewise, images falling inside the face space threshold would be identified as face images and subsequently evaluated against a gallery of stored face images for the closest match. Comparisons can be made in the lower dimensional face space using any of a variety of distance metrics aside from the Euclidean distance depicted above. These measurement options will be discussed later. An overview of the eigenface method is provided in Figure 15, that summarizes each step as well as the two possible outcomes where the distance measure ε, identifies it as a non-face and the process stops, or else ε falls within the face space threshold and Ω comparisons proceed to an identity [20].

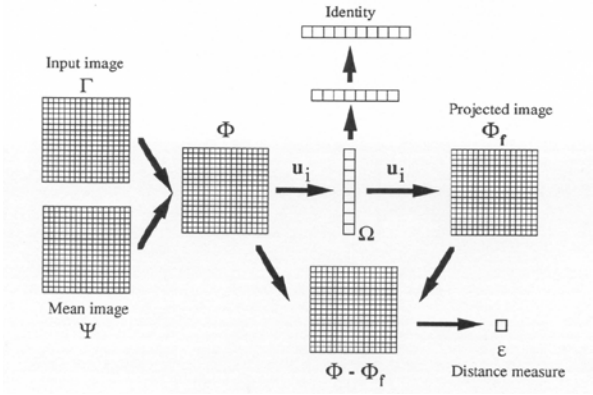

Figure 15: Summary of Eigenface Procedure [20]

Fisherface

The type of projection used to transform the original image into an alternative comparison space has a significant impact on the utility of the matching algorithm. The Eigenface algorithm was designed to exploit the holistic nature of the face image using PCA. Similarly, the Fisherface algorithm was designed to be insensitive to changes in lighting conditions and facial expressions using Linear Discriminant Analysis (LDA). These variations illustrated in Figure 16 can be some of the most troublesome for face recognition algorithms as the measured difference caused by these variations are greater than the difference between dissimilar faces.

Figure 16: Same Person with Different Expression and Lighting [22]

Using Fisher's LDA projection method, Belhumeur chose a linear projection that was orthogonal to the within-class variations (lighting and expression for a single subject) while maintaining the between-class variance (difference between different subjects) [22]. Comparatively, the Eigenface method uses a projection direction that maximizes variance across all images and therefore is subject to errors due to lighting and expression. Duda best summarizes the difference in the following example. PCA focuses on the gross features that characterize O's and Q's but might not notice the tail of the Q [23]. Discriminant analysis seeks directions that are useful for discrimination whereas PCA looks for efficient representation [23].

For the Fisherface method, a projection is chosen that maximizes the ratio of the between scatter and within scatter. The images ($x_1, x_2, ...x_i$), mean images ($\mu_1, \mu_2, ...\mu_i$) of the subject class χ_i , and the number of images in subject class *i*, $|\chi_i|$ are used to determine the between and within scatter as shown in Equation 9.

$$S_B = \sum_{i=1}^{c} |\chi_i| (\mu_i - \mu)(\mu_i - \mu)^T$$

$$S_W = \sum_{i=1}^{c} \sum_{x_k \in \chi_i} (x_k - \mu_i)(x_k - \mu_i)^T$$

Equation 9: Between and Within Class Scatter [22]

Since the within-class scatter S_W is singular, Fisherface first uses PCA to reduce the dimensionality of the face space, before calculating scatter using the previous equations.

Algorithm Extension for Pose and Internal Features

Pentland extended his initial work on eigenfaces with an effort that explored the benefit of a view-based eigenspace [24]. A view-based method can be constructed by developing comparison spaces using only images with the same pose and repeating this process for all possible poses. Using a distance from face space (DFFS) measurement, previously referred to as the reconstruction error in the eigenface overview, the orientation of an image can be determined based on which view space projection results in the smallest DFFS [20].

Determining the view space of an image allows the high dimensional face space to be divided into distinct regions. Prince devised an alternative approach for more challenging problems using a pose dependent linear transform to represent all poses called Tied Factor Analysis [25]. The goal of this approach was to address the problem of matching face images that are taken at a different pose compared those that are contained in the gallery. Prince's solution used a linear mapping to take images from an observed image space to an ideal identity space where an individual's salient features do not vary with pose. Although both approaches have been tested on benchmark datasets, there is a computational simplicity that the view-based technique offers especially when it is applied to large populations.

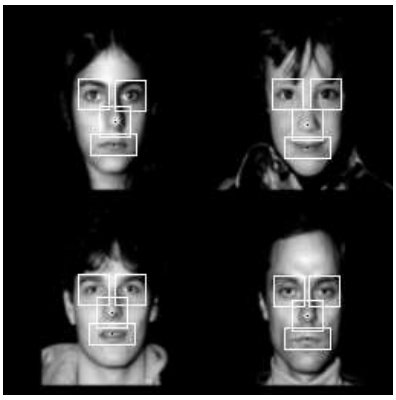

Figure 17: Eigeneyes, Eigennose, and Eigenmouth [24]

Pentland later modified the eigenface method to locate and identify facial features as depicted in Figure 17 [24]. This eigenfeature approach mitigates weaknesses of the earlier eigenface algorithm when faced with occlusions, disguises, and changes in expression. Insights from the cognitive research discussed earlier can guide the combination and weighting of prominent features according to their relative contribution to recognition.

Findings that acknowledged the importance of eyes and eyebrows recommend that these features be heavily weighted compared to other face features [4]. On the other hand, the mouth feature can be lightly weighted due to the likely variations in expression that can present a significant departure from recognizing a neutral expression. A confirmation of this shortcoming has been made by several state transportation departments as they have mandated that all new driver license photographs be taken with a neutral expression to aid their law enforcement identification systems. For occluded images, individual features that are present or visible would be assessed a larger weight

than the overall face image. A few illustrations of Pentland's successful extension and application of the eigenfeature method are depicted in Figure 18 [24].

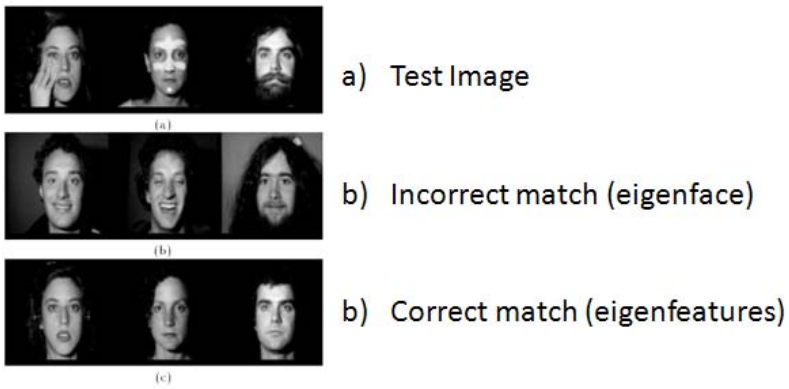

Figure 18: Eigenfeature Testing [24]

A combination of both the eigenface and eigenfeature methods into a collective score is shown to garner the highest performance as illustrated in Figure 19 [24]. Cognitive researchers point to a possible parallel in human recognition processing that starts with a low-resolution holistic representation for general identification supplemented by high-resolution facial features to verify our initial recognition assumption [24].

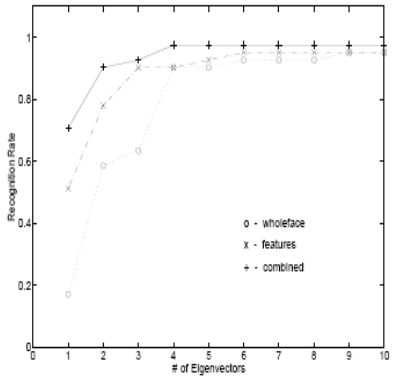

Figure 19: Recognition Rate for Eigenfaces, Eigenfeatures and Combined [24]

Neural Networks

A promising approach for complex pattern recognition is the application of neural networks (NN). Given the dimensionality of the face recognition problem and the desire to recreate the human cognitive ability, it is no surprise that NNs are regularly considered in face recognition applications. Duda replicated the previous data analysis technique of PCA with a three layer neural network with linear hidden units [23]. Although not discussed here, applicable techniques such as Nonlinear Component Analysis (NLCA) are also demonstrated by Duda in the form of a five layer neural network with two layers of sigmodal units and even the successful Independent Component Analysis (ICA) technique is demonstrated in a NN form [23].

Radial Basis Function NNs (RBFNN) are advantageous because they are universal approximators, possess the best approximation property, and display a compact topology with fast learning speeds [27]. A typical small training set for each individual and the accompanied high dimensionality of the image presents hurdles that these classifiers must first overcome. Some of the common difficulties involve over fitting, overtraining, and small sample effect may be addressed in preprocessing or in an adaptive NN application.

The general goal of NN design for face recognition is to develop a compact representation of faces capturing the features critical for identification. Pan [26] used one such approach in the application of NNs to classify individuals after the redundancy of important facial features were reduced using Discrete Cosine Transforms (DCT). Using only the DCT coefficients and the use of a back propagation NN a capable and fast

recognition system was developed using only grayscale images [26]. This successful compact transform displayed the ability to reduce data redundancy of the image but still maintain critical features for representing the mouth, eyes and hair [26].

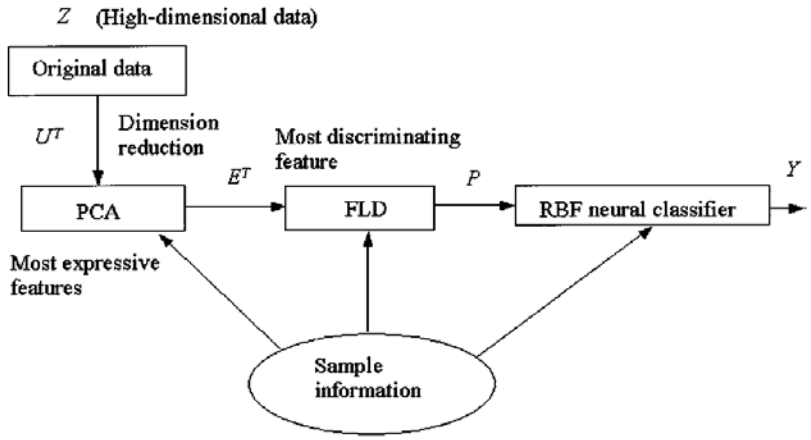

Figure 20: Diagram of RBF Neural Classifier [27]

Another example of a NN approach was used by Er [27] and is depicted in Figure 20. This application applied sequential techniques that enabled the RBFNN to take advantage of the respective strengths of PCA and Fisher's Linear Discriminant (FLD) method. Through PCA, the dimension of the face vector can be reduced to a number of dimensions that is more manageable. These features containing the highest variability are most useful for describing the data but not necessarily the best features for discrimination. Consequently, in the next step, FLD technique is used to identify the most useful features. The application of FLD to the projected training data can determine the best subspace for classification and separate the training data by maximizing the between-class scatter and minimizing the within-class scatter.

As mentioned earlier in the Eigenface and Fisherface applications, the solution to addressing particular shortcomings is usually a hybrid or fusion of complementary approaches. An assortment of algorithms can be used in combination on various feature sets to overcome common face recognition challenges. NN applications lend themselves to a single network approaches or just as easily to an ensemble of networks. Hansen [28] proposed the implementation of NN ensembles to first classify inputs and then to form a consensus on the output in order to reduce the overall generalization error.

The sampling of algorithms discussed are common examples to illustrate previous methods and their underlying logic. For a more thorough discussion of the numerous face recognition algorithms available, several very good literature surveys are available that include Abate [29], Samal [110], Kong [18], Zou [30], and Zhao [31]. Transitioning from academic research to commercial applications, the reoccurring vendor tests conducted by the National Institute of Standards and Technology (NIST) provide an excellent forum to illustrate the performance capability of commercial products but little insight into the guarded secrets of these proprietary applications [15][32]. The importance of algorithm design causes it to be both the central focus of research and then the heavily protected intellectual property of the technology industry. With a basic understanding of both the strengths and weakness of basic techniques, we move to other design aspects that need to be considered in the development of these systems. One important contributor to recognition performance is the construction of the comparison space where feature matching occurs.

Face Space

Various research efforts have examined the creation and dimensionality of the comparison space commonly referred to as the "face space." The gallery is comprised of all stored images in memory and generally, a subset of these gallery images serves as the training set. The face space is built using the training set and creates a smaller dimensional space comprised of uncorrelated features. This space should balance the requirement to represent the variety of faces and their unique features but with a dimensionality that allows the system to be computationally efficient. Academic research efforts often overlook this aspect, using closed experiments where all subjects contained in the gallery are used in the training set. When evaluating larger datasets, the point of diminishing returns becomes evident as the addition of training images no longer result in performance improvement.

In the eigenface approach, the effectiveness of the classifier can depend on the training set selected to create the face space. Chawla investigated the creation of face space by observing the performance of a random ensemble of random face spaces compared to a chosen face space obtained by selectively dropping specific eigenvectors [33][34]. Chawla also explored the benefit of combining multiple classifiers operating on an ensemble of subspaces with randomly chosen dimensions [33]. His results indicated that an ensemble of classifiers working on random subspaces is often competitive and sometimes better than a single classifier working on a select face space.

In a subsequent effort, Chawla created an algorithm that develops a face space based on the most diverse faces in a gallery [34]. The results obtained using a diverse

face space was more robust than a face space created from images with common illumination effects. The diverse set of images achieved desired performance levels with a reduced number of training images and subsequently smaller dimensional face space. When creating a training set it is advantageous to include an assortment of environmental effects and images to create a robust capability that can be employed outside the laboratory.

Once the selection of training images is determined, a subsequent decision must be made regarding the number of eigenfaces to retain. It is common practice to reduce the dimensionality of the face space by discarding eigenvectors corresponding to the highest and lowest eigenvalues. Eigenvectors associated with the highest eigenvalues are usually associated with the variation of illumination effects. Eigenvectors associated with the lowest eigenvalues are often associated with high frequency information that does not always contribute to improved recognition. This phenomenon was apparent in earlier cognitive examples as well. General rules of thumb for choosing the number of eigenfaces are recurring but established guidelines do not formally exist [33].

An initial trial and error approach is also commonly used to explore the performance limitations for a particular data set with their subsequent findings guiding the eventual implementation. Robinson [21] explored this aspect in a number of experiments and identified the diminishing level of performance of additional eigenfaces. The results of his experiments on three popular face databases (Rice, AT&T, and Yale) are shown in Figure 21 [21]. His findings reinforce the common practice of using a range of 5 to 10 prominent eigenfaces, as performance increases beyond this number experience diminishing returns.

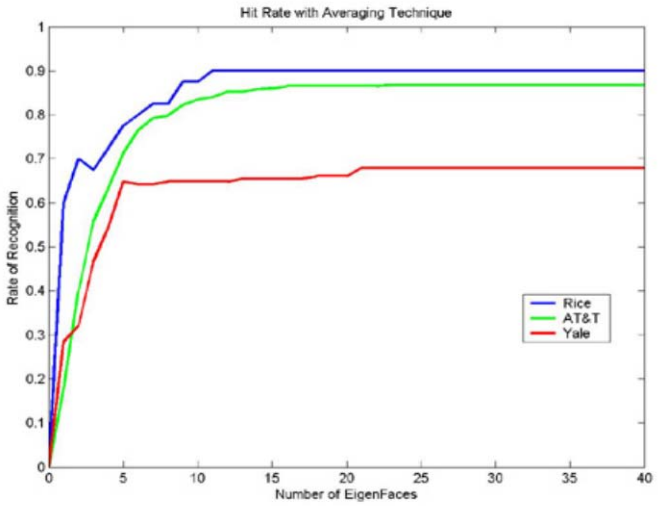

Figure 21: Recognition Capability with Increasing Number of Eigenfaces [21]

Contrary to this practice, Penev showed in his research that the dimensionality of the face space could require anywhere from 400 to 700 eigenfaces to adequately capture the fine detail necessary for larger and more diverse face databases [35]. His findings suggest that even general characteristics such as gender, ethnicity, expression, overall shape, and illumination require in the range of 200 eigenfaces. Based on his findings, Penev cautions on using a face space with too small of a face space, especially if the methods are applied to large datasets, high-resolution images, or those with additional variations such as pose or outward changes in appearance (i.e. glasses or hairstyles).

If this practice is employed on databases of ever-increasing size, the computational cost of this method could become prohibitive. For the widely used Eigenface algorithm, there are various opinions on the best implementation to represent the face space. Despite the number of research efforts focused on this consideration, there is still an assortment of opinions concerning this important consideration. A

mixture of selection practices seems to offer their own unique benefits but sometimes algorithm performance is more dependent on the data analyzed or the metrics used to measure similarity.

Distance Metrics

The selection of the distance metric is another aspect that can have a significant impact on the performance of the face-matching algorithm. Similar to considerations for the development of the face space, there are also a number of common distance metrics that have evolved to maximize recognition performance. A distance metric can be defined in the following manner (Equation 10).

$$
\begin{aligned}
&i.\ D(u,\ v) \geq 0 \\
&ii.\ D(u,\ v) = D(v,\ u) \\
&iii.\ D(u,\ v) \leq D(u,\ w) + D(w,\ v) \\
&iv.\ D(u,\ u) = 0
\end{aligned}
$$

Equation 10: Definition of Distance Metric

An array of these common distance measurements and calculations will be reviewed next. Two of the more familiar metrics used are the city block (Manhattan) and Euclidean metrics shown in Equation 11.

$$D_{CityBlock}(u,v) = \sum_{i=1}^{n} \left| u_i - v_i \right|$$

$$D_{Euclidean}(u,v) = \sqrt{\sum_{i=1}^{n} \left(u_i - v_i \right)^2}$$

Equation 11: Common Distance Metrics

Usually distances are measured and used in a manner where a smaller distance represents a closer match. In the case of the correlation metric (Equation 12), the metric range of -1 (negative correlation) to 1 (positive correlation) is mapped into a range of 0 to 2 for consistency. Likewise, the covariance metric (Equation 13) that produces a 1, when the vectors are the same, and a zero, when they are orthogonal, is subsequently flipped to produce a 0 to 1 distance measurement.

$$S_{Correlation}(u,v) = \frac{\sum_{i=1}^{n} \left(u_i - \overline{u} \right) \left(v_i - \overline{v} \right)}{(N-1) \sqrt{\frac{\sum_{i=1}^{n} \left(u_i - \overline{u} \right)^2}{N-1}} \sqrt{\frac{\sum_{i=1}^{n} \left(v_i - \overline{v} \right)^2}{N-1}}}$$

$$D_{Correlation}(u,v) = 1 - S_{Correlation}(u,v)$$

Equation 12: Correlation Distance Measurement [36]

$$S_{Covariance}(u,v) = \frac{\sum_{i=1}^{n} u_i v_i}{\sqrt{\sum_{i=1}^{n} u_i^{\,2}} \sqrt{\sum_{i=1}^{n} v_i^{\,2}}}$$

$$D_{Covariance}(u,v) = 1 - S_{Covariance}(u,v)$$

Equation 13: Covariance Distance Measurements [36]

As highlighted and tested by Miller [37], a distinction between the vectors representing the faces being compared needs to be differentiated when we are talking about a PCA or LDA space or a Mahalanobis space. In PCA and LDA spaces, the respective face vectors being compared have a sample variance, σ_i^2 , that is represented by their eigenvalues, λ_i, whereas in Mahalanobis space the variance is one. Using the formulation by Beveridge [36] and Miller [37], the Mahalanobis distance metrics, as u_i and v_i represent the associated face vectors in PCA/LDA space and m_i and n_i in Mahalanobis space as shown in Equation 14.

$$m_i = \frac{u_i}{\sigma_i} \qquad n_i = \frac{v_i}{\sigma_i}$$

Equation 14: Relation between PCA/LDA and Mahalanobis Space

The city block and generic Mahalanobis distance are similar to their counterparts in PCA space but simply adjusted to Mahalanobis space. These and other Mahalanobis distances are illustrated in Equation 15 [36].

$$D_{MahalanobisCityBlock}\ (u,v) = \sum_{i=1}^{n} \left| m_i - n_i \right|$$

$$D_{Mahalanobis}\ (u,v) = \sqrt{\sum_{i=1}^{n} \left(m_i - n_i \right)^2}$$

$$S_{MahalanobisCosine}(u,v) = \cos(\theta_{mn}) = \frac{|m||n|\cos(\theta_{mn})}{|m||n|} = \frac{m \cdot n}{|m||n|}$$

$$D_{MahalanobisCosine}(u,v) = -S_{MahalanobisCosine}(u,v)$$

Equation 15: Mahalanobis Distance Measurements [36]

The Mahalanobis Cosine is the cosine of the angle between the two face vectors and is adjusted for a distance-like measurement. While the Mahalanobis Cosine distance measurement has gained recent popularity and is often the default metric used in testing, some additional distance measurements, shown in Equation 16, have shown results comparable to the more traditional metrics.

Research in this area has found that certain metric and algorithmic combinations perform better together but the overall success can also be affected by the database being used or even the individual face image that is being evaluated [37]. Given the importance of this selection, common testing environments such as the Colorado State University's Face Identification Evaluation System [36], for example, have continued to incorporate a number of different metrics over time to investigate their algorithms and databases of grayscale face images. Some additional distance measurements explored in the literature are shown in Equation 16.

$$D_{Hellinger}\ (u,v) = \sqrt{\sum_{i=1}^{n}\left(\sqrt{|u_i|} - \sqrt{|v_i|}\right)^2}$$

$$D_{Canberra}\ (u,v) = \sqrt{\sum_{i=1}^{n}\frac{|u_i - v_i|}{|u_i + v_i|}}$$

$$D_{Czekanowski}\ (u,v) = \frac{2 * \sum_{i=1}^{n}\min(u_i, v_i)}{\sum_{i=1}^{n} u_i + v_i}$$

Equation 16: Additional Distance Measurements [37]

Biometric Features

A majority of face recognition applications utilize the overall appearance and spatial features of the face. However, there are additional features that can be exploited to contribute to identify faces. The role of soft biometrics, spectral signatures, and internal features are discussed in the following section.

Soft Biometrics

A face image contains traits that indicate gender, age, and ethnicity as well as unique supplemental characteristics related to their skin and hair that are valuable clues to describe and identify an individual. Jain [38] defines soft biometrics as characteristics such as these "that provide some information about the individual, but lack the distinctiveness and permanence to sufficiently differentiate any two individuals." In

Jain's fingerprint research [38], he demonstrated that the use of soft biometric information can significantly improve fingerprint recognition performance. The same type of enhancement should be possible in an application to face recognition.

Gender extraction has been performed by Balci [39], who used a multilayer perceptron (MLP) for gender classification in order to divide the face space into a male and female population. Impressively, Moghaddam [40] was able to successfully implement support vector machines to characterize gender from as little as a low-resolution thumbnail image taken from the popular Face Recognition Technology (FERET) database [41]. Another relevant factor related to gender is the fact that in commercial testing the identification rate for males is generally higher (6-9%) than females [32].

Age is another soft biometric category that can be extracted from face images and has been demonstrated in research experiments. Kwon [42] devised a method to classify images as an infant, adult, or senior based on ratios calculated from the location of prominent facial features and the presence of wrinkles on the forehead and below each eye. The underlying theory is based on the fact that children's facial features lie predominantly in the lower portion of the face and that they gradually shift towards the middle of the face over time as the jaw and lower portion of face grows. Older adults are distinguishable because of the increased number of wrinkles that comes with age.

Horng [43] extended this work by classifying images in 4 separate age categories. Using sobel edge detectors, he located prominent facial features (eyes, nose, and mouth) and then determined wrinkle features such as density, depth and skin variance. The

relative geometry of prominent face features is fed into the first of two back propagation NNs to identify infants [43]. Three wrinkle features of density, depth, and skin variances, are used in the second NN to further divide the non-infants into young, middle age, and senior adults.

Mukaida [44] offered an alternative to applying global filters to identify wrinkles and spots from a grayscale image. Using binary images and assessing the size, shape, and density of image blobs, Mukaida differentiates between primary facial features and secondary characteristics of wrinkles and spots [44]. By exaggerating or diminishing wrinkles and spots, he was able to change the appearance and estimated age of an individual [44]. Mukaida's results suggest that humans estimate age based on spots, wrinkles, and face parts individually and then integrate them for an overall age assessment. The ability to distinguish the wrinkle and spot features could prove very beneficial as distinguishing marks for older adults. This soft characteristic of age has also been identified as a factor affecting the accuracy of face recognition systems where the older an individual is the easier it is to identify them. In commercial testing it was found that recognition rates increase about 5% for every 10 years of age up through age 63 [32].

Gutta [45] looked at both gender and ethnicity through the use of RBFNN and inductive decision trees. The results were impressive for both gender and ethnic classification with relative success rates of 94% and 96%, respectively. Gutta proposes that a similar approach be utilized to discriminate images based on gender, age, and ethnicity before attempting recognition [45].

With the ability to distinguish these descriptive characteristics automatically, the incorporation of soft biometrics can be applied in the same manner as it was in fingerprint recognition to enhance face recognition. In one example, Marcel [46] took a grayscale image and combined it with the RGB (Red, Green, and Blue wavelength) color distribution vector to improve overall recognition performance. His results are in line with the cognitive research findings discussed earlier. These results indicated that there is enhanced value of pigmentation or skin color in recognition applications [46], [4].

The ability to use these soft biometric traits to enhance face recognition performance has been limited to single traits only and predominantly in the visible portion of the electromagnetic spectrum. The sum of extracted soft features such as age, ethnicity, gender, and proportion of genuine and visible face surface (i.e. not masked with make-up, accessories or occluded) can provide valuable context in which to make identification decisions. This awareness also offers an ability to select individuals that may be trying to avoid detection and warrant scrutiny.

An enormous efficiency can be obtained, if a population of subjects can be quickly narrowed based on soft characteristics since only a fraction of the total images would have to be processed for matching. Phillips [32] showed in the Face Recognition Vendor Tests that the accuracy of our best commercial systems is reduced approximately 2-3% as the size of the database doubles in size. By utilizing and interpreting contextual information contained within face imagery, it is possible to partition the number of faces in the gallery resulting in improved performance and efficiency.

Spectral Properties

Face recognition applications continue to be challenged by variations in illumination. The image that is captured by the eye or camera is dependent on the direction, amount, and type of illumination as well as the optical properties of the human skin. The spectral reflectance of skin is mainly determined by the presence of pigment, or melanin, and blood, specifically oxygenated hemoglobin [47]. Human skin is comprised of several layers starting with a very thin outer layer of the stratum corneum, followed by the epidermis and then the underlying dermis as depicted in Figure 22 [48].

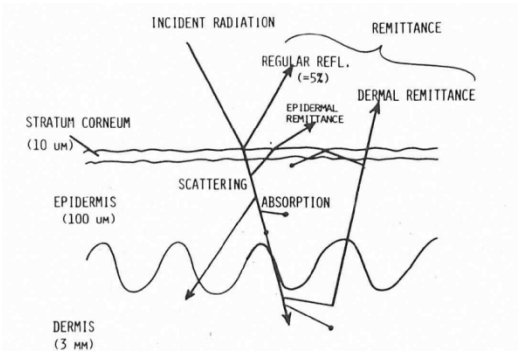

Figure 22: Optical Properties of the Skin [48]

Melanin is found both in the top two layers, epidermis and stratum corneum, while hemoglobin is contained in the dermis. Different wavelengths of the electromagnetic spectrum are able to penetrate human tissue at different depths. Therefore, the reflectance properties of skin are both wavelength and tissue dependent [48]. When these two factors are examined together throughout the visible and infrared (IR) spectrum, discernible characteristics become evident that can be used to locate, segment and characterize faces.

Spectral remittance is the amount of light returned by a surface by reflection or scattering. The range of differences in the spectral remittance is depicted as the spectral signature of both light and dark skin is shown in Figure 23. Melanin is the main absorber of radiation in the spectrum from 350-1200 nm [48]. As illustrated in Figure 23, absorption decreases for longer wavelengths resulting in a higher remittance up until 1200 nm [48]. Beyond this range skin remittance is unaffected by melanin content [48].

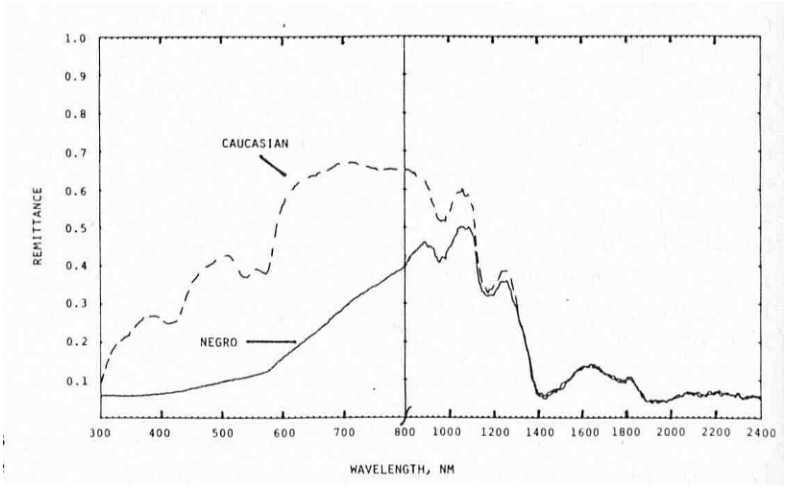

Figure 23: Variations in Spectral Remittance of Skin [48]

Higher amounts of melanin decrease the remittance and relative slope of the spectral signature of human skin. The presence of hemoglobin can be seen in the absorption bands at 410 nm, 540 nm, and 575 nm, represented by the resulting dips in the light skin signature in Figure 23 [48]. This effect is masked by the increased absorption of the melanin in dark skinned signatures [48]. Storring [49] demonstrated that by modeling these reflective characteristics in the visible wavelengths it was possible to detect human skin under changing and mixed lighting conditions.

Melanin serves to protect humans from harmful ultraviolet radiation of the sun and over time a population's melanin level adjusts to the environment. Studies have shown that melanin levels of indigenous people are strongly correlated to their absolute latitude on the earth and the levels of ultraviolet (UV) radiation exposure. Spectral skin signatures are highly dependent on melanin levels and can be extended to determine a range of ethnicities. The mean spectral reflectance for different ethnicities are depicted below in Figure 24 [47].

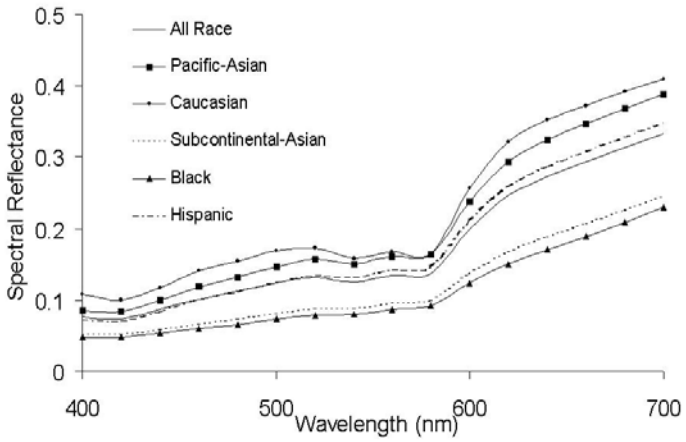

Figure 24: Differences in Mean Spectral Reflectances [47]

Melanin levels may also offer clues as to the gender of a particular individual. In populations analyzed by Jablonski [50], it was found that females typically had lighter skin pigmentation than males. This could be biologically attributed to the female's need for increased vitamin D production during pregnancy. Among homogeneous populations, this finding may be able to assist in gender determination. Jablonski used a linear regression of UV radiation measurements worldwide and skin spectrometer

readings from numerous indigenous populations to create a predicted shading of skin colors around the world [50]. The results are shown in Figure 25 [50].

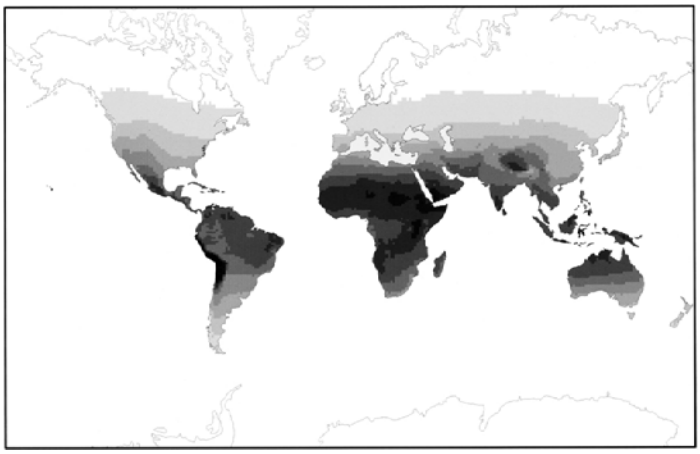

Figure 25: Predicted Spectral Reflectance based on UV Exposure [49]

It is interesting to note that a previously believed limit to face recognition performance may be affected by the characteristics of the skin and its reaction to environmental factors. Daugman, the biometric pioneer and computer vision expert from Cambridge University, known for creating the iris recognition algorithm used in all commercial systems, noted that an upper bound on the performance of face recognition is limited by the birth rate of identical twins [51]. However, at the cutaneous level there is an accumulation of empirical environmental data that shows itself in the development of nevi and freckles.

Bataille [52] in her research on risk factors for melanoma for twins, quantified the differences between monozygotic and dizygotic twins and noticed that for environmentally exposed surfaces, i.e. the face, there is an increasing amount of variance with increased sun exposure and age. These features are evaluated by the medical

community through Ultra Violet (UV) illumination of a Wood's Lamp to accentuate age spots and skin abnormalities in order to diagnose skin disorders. These spots and patterns can also serve as supplemental features for matching algorithms if needed.

Melanin, contained throughout a person's skin, is present in smaller quantities in the lips. Consequently, the hemoglobin absorption bands mentioned at 540 nm and 575 nm are more pronounced resulting in more pronounced dips in the remittance curve for lip tissue, represented by the green curve in Figure 26 [47]. This characteristic can be exploited to locate and track a person's lips for gesture recognition or incorporated into speech recognition applications.

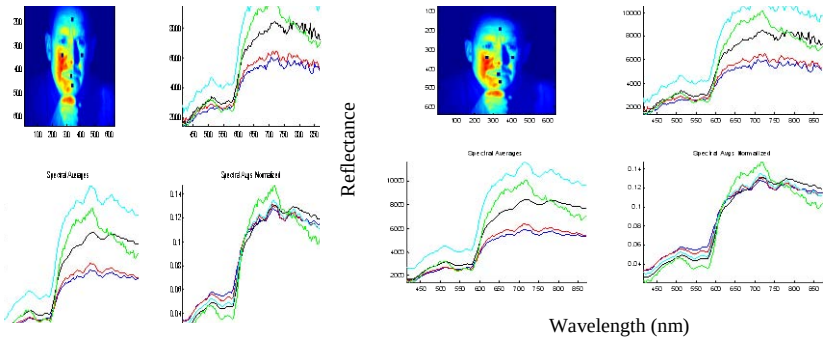

Figure 26: Spectral Signature of Lips (Green Line) Using Visible Wavelengths

Unique spectral characteristics are also present in the human eye or a subject's hair. The eye, or more specifically the iris, contains the pigment material melanin. With a different biological structure than skin, the remittance properties offer unique features that can map to the soft biometric of eye color. The utility of this type of approach was explored in a study by Boyce [53] using multispectral information from near IR (NIR) and visual wavelengths (RGB) in an iris recognition application. In this application the

use of multispectral imagery showed the potential to improve segmentation and the overall performance of iris recognition systems [53]. Color information was clustered to examine various components of the iris and various wavelengths matching scores were fused with a simple sum rule for increased accuracy [53].

In human hair, melanin is also present in the form of eumelanin (brown-black) and phaemelanin (yellow-red) pigments [47]. When large amounts of melanin are present the hair appears darker. For smaller amounts of melanin, the colorless outer skin tends to dominate, making the overall appearance of the hair lighter. As individuals age, hair tends to lose melanin altogether giving it the white or gray appearance [47]. The unique signature of the hair can help locate and segment important portions of the face including hairlines, eyebrows, beards and mustaches. This can also serve to highlight inconsistencies as individuals try to alter their appearance with hair extensions, dyes, or wigs. Pavlidis [54] illustrated the ability to detect wigs and toupees as individuals attempted to alter their appearance by using upper-NIR wavelengths (1.3 - 1.7 μm).

The extraction of the soft biometric of hair color can provide useful clues that can be used in a facial recognition application. An example of hair signature compared with skin signatures is shown in Figure 27. Nunez [54] extended this research by characterizing a person's hair color based on its spectral signature as depicted in Figure 28. There are some useful characteristics captured in the wavelengths beyond 1200nm, where melanin induced remittance becomes less significant, as mentioned earlier in Pavlidis' findings.

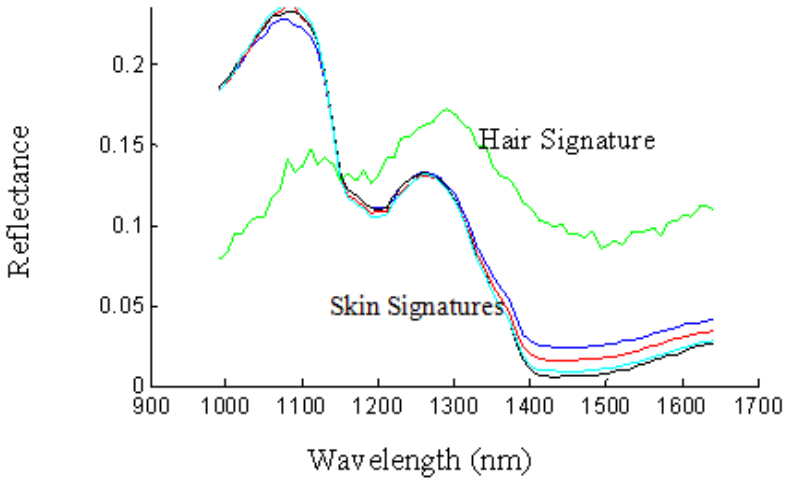

Figure 27: NIR Spectral Signature of Hair vs. Skin

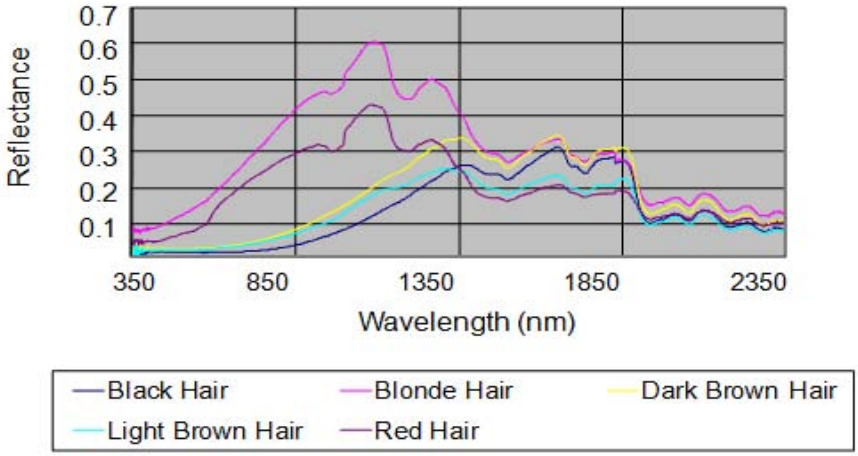

Figure 28: Spectral Signatures of Various Hair Colors [54]

The ability to identify hair spectrally can also be applied to the important feature of eyebrows. The eyebrows are central to human recognition abilities as their position and movement are often correlated with specific emotions. With their stark contrast to the surrounding skin segments of the face and their location between the flat forehead and the eye orbit, the eyebrows maintain a prominent position in the center of the face on a

peak between the forehead and eyes. These characteristics make them an important light and shadow resistant feature.

The importance of this feature was not overlooked in the previously mentioned work of Sinha [4]. His findings show that the ability of humans to recognize faces without eyebrows is more negatively affected that images with the eyes removed all together [4].

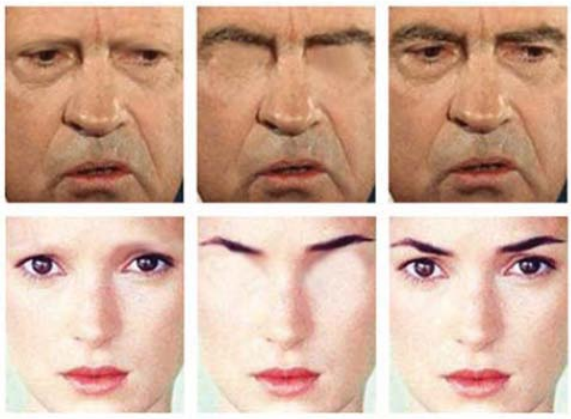

Figure 29: Face Representations without Eyes and Eyebrows [4]

In Figure 29, images of former President Nixon and celebrity Winona Ryder are shown without eyebrows, eyes and normally for comparison [4]. The eyebrows are a feature that can be critical in either a feature-based identification or a semantic application focused on emotion or behavioral recognition.

Face images contain an impressive amount of information from soft biometric features to spectral signatures. This data can be utilized in a number of ways to enhance the performance of existing face recognition systems. It has been shown that simple soft biometric indicators can improve the performance of other biometric systems. Several

researchers have made the automatic extraction of these features such as gender, age and ethnicity from an ordinary image possible. With the added dimensions of HSI, fiducial features of the face can be easily identified, segmented, and categorized. Combined together the mutual information and diversity of these features can provide valuable assistance for improving face recognition systems.

Recognition Challenges

Despite the advancement of face recognition technology, challenges remain in the reoccurring areas of occlusion, illumination, pose, expression, and circumvention. Progress has been made on all of these fronts, but no single approach or algorithm has provided a solution to these problems. In the following section a range of techniques from the literature are discussed as a prelude to implementing them in a hyperspectral face recognition system.

Occlusion

Although capturing face images is not an intrusive act, it is relatively easy for a non-cooperative subject to cover a portion of their face with a scarf, sunglasses or hand in an attempt to avoid identification. Robinson [21] explored the effect that occluded images, as well as blurred images, have on face recognition using the eigenface method. Vertical and horizontal occlusions of various widths were applied to the face images in a manner depicted in Figure 30.

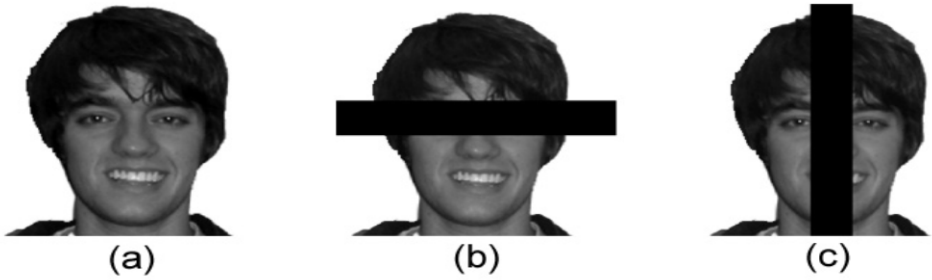

Figure 30: Face Images (Undistorted, horizontal, and vertical occlusions) [21]

The holistic eigenface method demonstrated a robust performance across many of these experiments, but predictably showed a degradation in performance as the size of the occlusion increased (1, 10 and 40 pixels). The results for horizontal and vertical obscured images are shown in Figure 31.

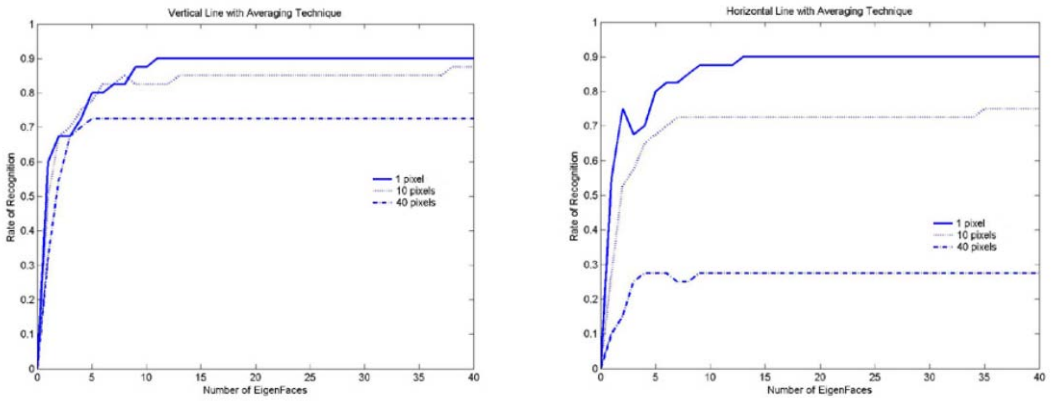

Figure 31: Recognition Rates with Increasing Occlusion Size [21]

Similarly, experiments with blurring were conducted using increasing level of blurring with a two-dimensional boxcar blur of various sizes (2, 20, and 40 pixels). The effect of the blurring is depicted in Figure 32.

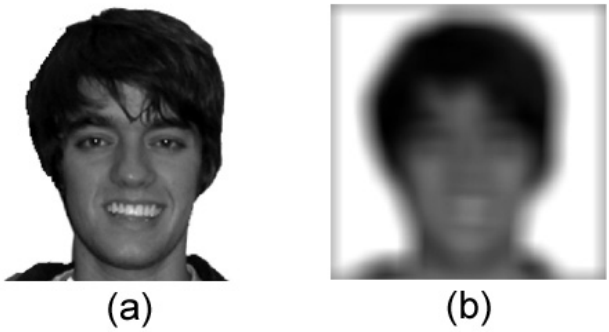

Figure 32: Original and Blurred Images [21]

This type of resolution can be common in poor lighting situations or at increased distances with common security cameras. Again, the results displayed decreasing performance as the blurring effect was increased (Figure 33) [21]. During testing the diminishing returns of performance is evident as the rate of recognition plateaus relatively quickly as additional eigenfaces are added

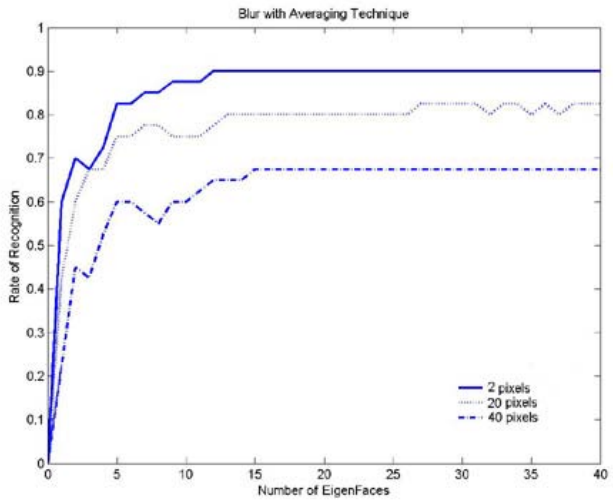

Figure 33: Recognition Rates for Blurred Images [21]

An interesting approach used by Robinson [21] was a method of averaging face weights for all images of a particular individual. The common technique used in many applications is to project and use all available images in the gallery for matching, realizing that most individuals will have several different pictures in the gallery. The averaging technique takes all face weights for a particular individual and averages them together to create a face class for that person. This practice showed increased performance rates in testing on the occluded and blurred images. This approach to create a face class could be replicated with images from different wavelengths or from different still images taken from a video stream.

Jenkins [56] also demonstrated increased performance in averaging faces together in a test of celebrity images. In this effort all images of an individual were combined to create an average texture model and average shape model. These models were then combined to improve a true positive rate from 54% to 100% for over 400 photographs of male celebrities. Jenkins [56] and Robinson's [21] results are reinforced by human recognition research by Sinha [4] that shows human recognition ability is higher for faces that are more familiar to us. Through these types of averaging processes environmental variations can be mitigated while creating more robust facial features.

Illumination

Illumination effects are a natural occurring variation that challenges all face recognition systems employed in an uncontrolled environment. These effects are difficult to overcome in the visible spectrum and become more complicated in a multispectral or

HSI. Current research, techniques and the advancement in sensor technology offers several options to overcome this challenge for greater performance.

To compensate for illumination within the spatial domain of the image, the previously discussed eigenface method can be modified. The imprecise practice of discarding the first few eigenvectors, presumably containing the largest variation of uncontrolled lighting, is commonly practiced to gain general improvement in performance.

When operating in the frequency domain of the image, Er [57] used a discrete cosine transform (DCT) to reduce the dimensionality of the data. A similar dimensionality reduction can be accomplished across large datasets quickly by using DCT and discarding low frequency coefficients. Er explored the effectiveness of this technique by focusing on the coefficients representing the illumination variation in images. In his approach, the first DCT coefficient represents the brightness of an image so discarding it or adjusting it to specified value can produce an image that is free of the major illumination variations and can be used more effectively for identification (Figure 34) [57].

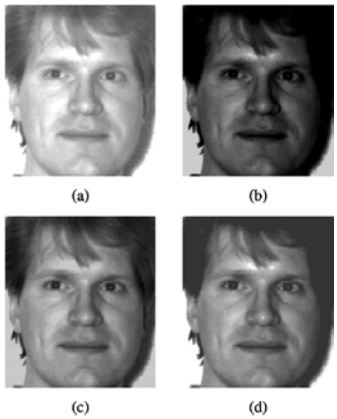

Figure 34: Adjusting Illumination with Discrete Cosine Transform [57]

In Figure 34, images (a) and (b) are the same image but taken at different levels of brightness. Images (c) and (d) are the adjusted images of (a) and (b) by using the technique of discarding the illumination component using DCT. This method can also be modified for asymmetric lighting effects where one side of the face is well illuminated while the other side is submerged in shadows. By measuring the average brightness for each face half and then correcting the imbalance, the image can be adjusted for more accurate comparisons. For more complicated non-uniform illumination variation, additional coefficients would have to be discarded or adjusted.

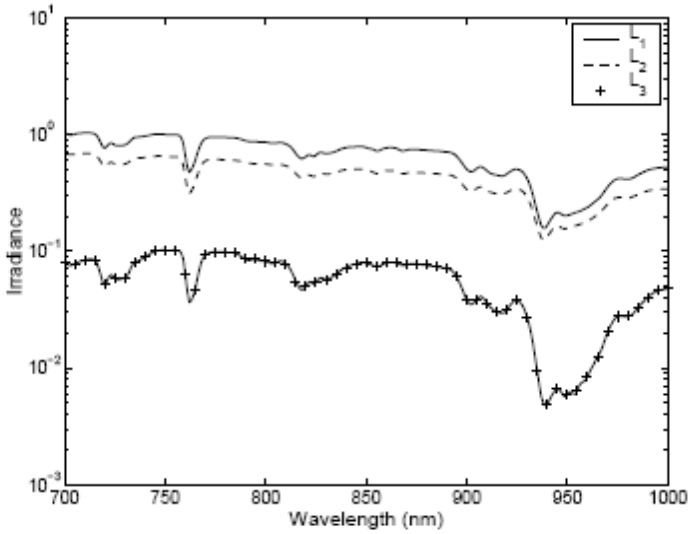

Figure 35: Global Irradiance Spectra Measured at Different Times [58]

Although Pan's [58] research is discussed later in this document, one of his initial investigations into hyperspectral face recognition was in an attempt to overcome illumination variations. A change in setting, location, or even the time of day can affect the entire spectrum of environmental illumination as natural lighting effects change during the day. A few samples of environmental lighting signatures are displayed in

Figure 35 to illustrate this variation. This can have a considerable effect on the resulting skin signatures, even signatures taken from the exact same tissue of the same person. Two examples taken from the forehead of the same person under different environmental illumination are displayed in Figure 36, showing that both the amplitude and shape are effected making any attempt at matching difficult [58]. In an effort to capture these effects, Pan collected thousands of outdoor illumination spectra to create a low dimensional linear model for illumination. Subsequent recognition was accomplished by first projecting face signatures into common illumination subspaces before comparison.

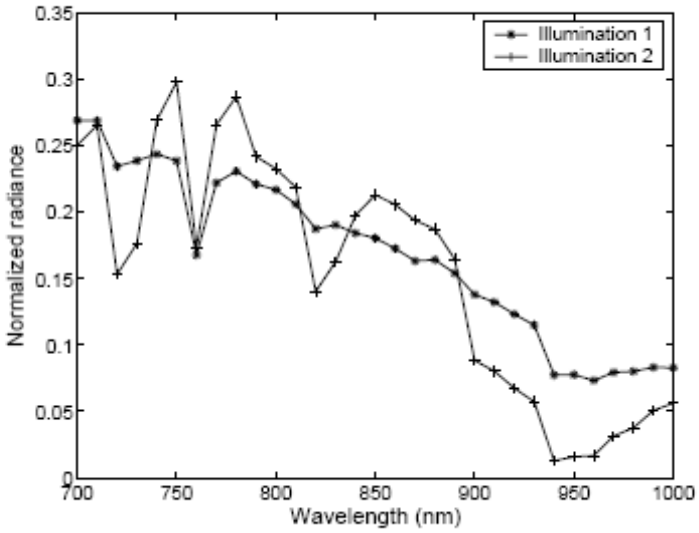

Figure 36: Spectral Signature of Forehead under Different Illumination [58]

One method that can be used to adjust for lighting conditions is to use a spectroradiometer when the image is initially taken. A spectroradiometer is used to measure ambient light conditions and is typically used to measure the spectral properties of lighting equipment. The spectroradiometer measures spectral irradiance (watts/m^2/nm) at contiguous wavelengths and if these measurements are sampled when the image was

taken, then they can be subtracted from the person's face image to obtain a 'ground truth' for HSI comparisons [59].

Pose

In addition to the eigenview [24] and tied factor analysis [25] mentioned earlier, there are other innovative approaches that can be used to compensate for pose variations. One notable method is Lowe's SIFT method [60] [61]. In this method, invariant features are found by searching for locations that are identifiable across variable scales, or throughout the scale space [61]. Scales in this discussion are different representations of the linear dimensions, much like you would see across maps of differing size or scope. Fortunately, on maps the scale format used is defined in a easily identifiable legend. For images, this information is not always so obvious. For that reason, invariant scale space features are used that are also robust to image translation and rotation as well as having some invariance to illumination changes and obscuration. A brief review of this method follows.

These interest points (x, y) are identified in the scale function, $L(x, y, \sigma)$, of the image and are located by convolving the image, $I(x, y)$, with a Gaussian function, $G(x, y, \sigma)$ [61].

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

Equation 17: Scale Function [60]

The familiar Gaussian function used in Equation 17 is defined below.

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

Equation 18: Variable Scale Gaussian Function [60]

A difference of Gaussian (DoG) filter is developed by using variable scales, *k*, of a constant multiplicative factor [62]. In this process, a blurred version of an image is subtracted from the original image removing certain spatial frequencies acting much like a band-pass filter to sharpen detailed information or edges in the image.

$$f(x, y, k, \sigma) = G(x, y, k\sigma) - G(x, y, \sigma)$$

Equation 19: Difference of Gaussian Filter [62]

The convolution of the Gaussian function with the image is conducted by the difference of two nearby scales [61].

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$$

$$= L(x, y, k\sigma) - L(x, y, \sigma)$$

Equation 20: Difference of Two Nearby Scales [61]

The images obtained are grouped by octave, where an octave is represented by the doubling of the σ value. These images are then subtracted from their adjacent or neighboring image to give us DoG images. Afterwards, the original scaled image is down-sampled by a factor of two and then the subtraction process is repeated (Figure 37).

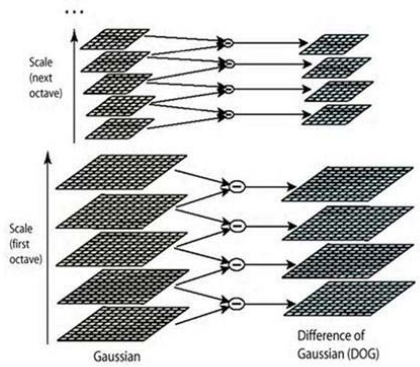

Figure 37: Producing the Set of Scaled Images [61]

Each pixel in the DoG images is compared with its immediate neighboring pixels as well as its bordering pixels in the DoG image at an adjacent scale. If the pixel is identified as a local maximum or minimum, with respect to the intensity, then it is recognized as a candidate key point. This process is illustrated in Figure 38 .

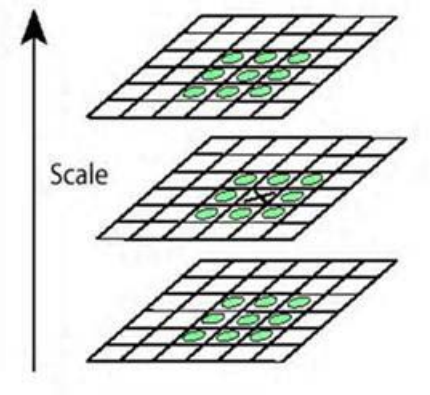

Figure 38: Minima and Maxima of the DoG Images [61]

Candidate key points are evaluated and points with low contrast that may be susceptible to noise are removed. In addition, candidate key points that occur along edges that are poorly located for matching can also be eliminated. For the key points that

remain, an overall orientation is determined by evaluating the gradient of neighboring pixels. The gradient magnitude, *m(x, y)*, and orientation, *θ(x, y)*, are computed using pixel differences as shown in Equation 21.

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}$$

$$\theta(x,y) = \tan^{-1}((L(x,y+1) - L(x,y-1)) / (L(x+1,y) - L(x-1,y)))$$

Equation 21: Gradient Magnitude and Orientation Calculation [61]

The gradient of neighboring pixels is measured and used to create an orientation histogram where the predominant orientation serves as a reference direction enabling a matching capability for rotated images.

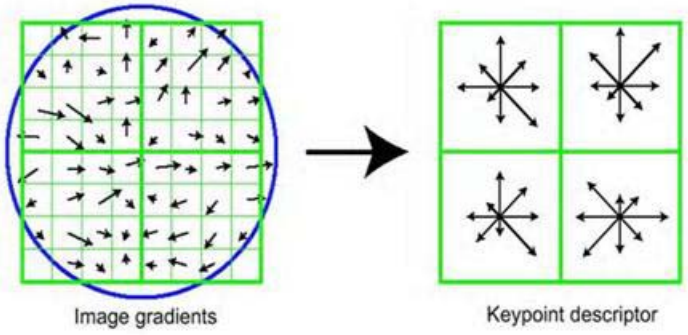

Figure 39: Creating the Keypoint Descriptor [61]

Using a group of orientations, a feature vector or descriptor is assembled that describes the local orientation and gradient [61]. This feature vector can be normalized to enhance the invariance to illumination effects. The feature vector is now ready for

matching using a variety of nearest neighbor methods. An illustration of these methods as they are applied to face images follows.

A demonstration of the SIFT method will be applied to a picture of Margaret Thatcher, the former Prime Minister of the United Kingdom, and a common example used to illustrate the previously discussed Thatcher illusion. In Figure 40, both the key points and associated reference vectors are displayed on the image.

Figure 40: Scale Invariant Feature Transform (SIFT) - Key Point Detection

The technique is applied to both an upright an inverted image and the displayed matching points are shown in Figure 41. In Figure 42, a different image, this time using the well known politician Secretary Clinton, an example is created to demonstrate the algorithms ability to overcome changes in both scaling and rotation as the matches are shown on images with varying size and orientation.

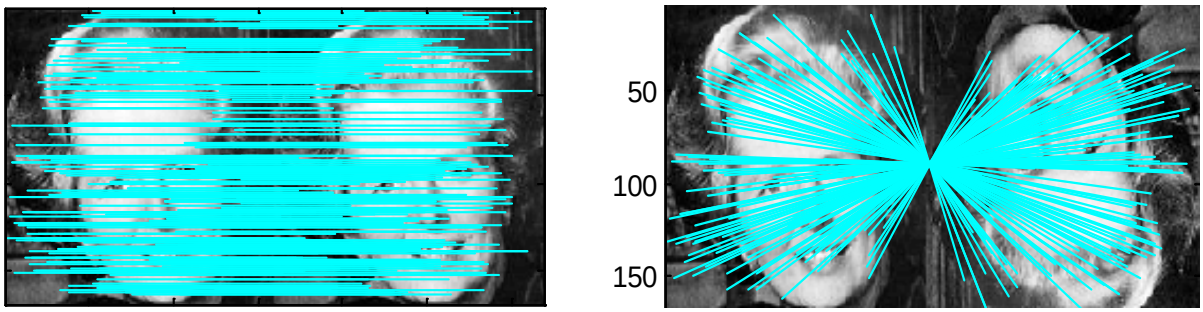

Figure 41: SIFT Applied to Upright and Inverted Images

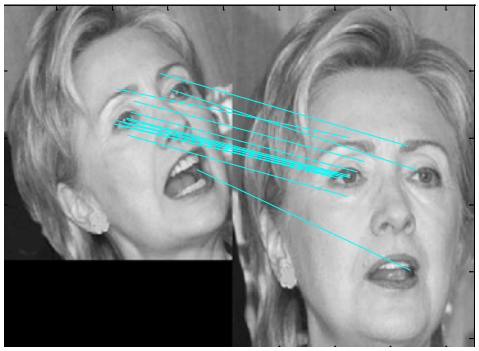

Figure 42: SIFT Applied to Scaled and Rotated Image

There have been a few research efforts that have explored the application of Lowe's SIFT method to face recognition. Mian [63] used 2D and 3D imagery to correct for pose variation. The interesting use of this SIFT application was the implementation in a rejection classifier to reduce the size of large galleries of potential candidates prior to matching [63]. Luo applied the SIFT method to face recognition by using a combination of both local and global similarity measures of SIFT features [64]. The local features derived from a K-mean partitioning of the face's SIFT features and a global representation of the entire face resulted in a method that performed well but still struggled when faced with variations in illumination and age [64].

Circumvention

In most face recognition research, the participants, eager to be good test subjects, willfully submit to having their faces being photographed. However, it has been shown that when individuals are evasive or deceptive, the face recognition performance quickly degrades. In one demonstration, Ramanathan [65] used an assortment of images with various poses and disguises to test the reliability of a recognition application (Figure 43).

Figure 43: National Geographic Face Images for Circumvention [65][66]

The application used implemented an eigenface algorithm that was combined with a half-face matching capability to help overcome illumination variances across the face. In the experiment, a trained CIA operative, using both disguises and photographs over different time periods of his life, easily confused the face recognition system to the point that the performance of the algorithm was no better than flipping a coin.

Singh [67] was able to improve on this performance with the goal of tackling both the challenge of identifying disguised faces but also the often associated difficulty of recognition with only a small sample of stored images for an individual. By using a 2D log polar Gabor transform within an artificial neural network architecture his approach showed improvements over many leading algorithms by capitalizing on capturing the textural features of the face in a manner that is similar to the way the human eye samples and processes images [67].

The challenge of non-cooperative subjects requires the use of additional measures to aid face recognition. One such technique may be by utilizing the unique spectral signatures from an individual's skin, eyes, and hair. Nunez [68] showed that HSI can offer the ability of a live skin test that utilizes the unique reflectance properties of both the skin's melanin and the oxygenated hemoglobin beneath it. In Figure 44, skin detection is demonstrated as the NIR spectral signature from a human face is compared to a similarly colored doll face.

Figure 44: NIR Skin Detection Compared to Doll's Face

The ability to use the signatures to detect skin can be used in a manner very similar to the Normalized Difference Vegetation Index (NDVI) used in remote sensing applications to detect live vegetation. Based on early work by Nunez [54], [68] a Normalized Differential Skin Index (NDSI) can be computed easily through the sum and difference of key spectral bands in the same way. In Nunez's work, he devised both a relative measure of a Normalized Differential Skin Index (NDSI, γ_i) and a Normalized Differential Melanin Index (NDMI, η_i) that proved useful for distinguishing different types of materials as well as a descriptive melanin measurement from an individual's

skin. These calculations are shown below in Equation 22, as $\hat{\rho}_i(\lambda)$ represents the reflectance of the i[th] pixel at wavelength λ [68].

$$\text{NDSI} \quad \gamma_i = \frac{\hat{\rho}_i(1080\text{nm}) - \hat{\rho}_i(1580\text{nm})}{\hat{\rho}_i(1080\text{nm}) + \hat{\rho}_i(1580\text{nm})}$$

$$\text{NIMI} \quad \eta_i = \frac{\hat{\rho}_i(1080\text{nm}) - \hat{\rho}_i(1180\text{nm})}{\hat{\rho}_i(1080\text{nm}) + \hat{\rho}_i(1180\text{nm})}$$

Equation 22: NDSI and NIMI Calculation [68]

From the simple NDSI calculation, a quick association could be made between the NDSI value and various materials that routinely act as confusers when attempting to utilize only the color characteristics of human skin for segmentation [68]. The larger water content of human skin results in NDSI values that are higher than common materials found in the background of an outdoor environment. A list of these materials and relative values are displayed in Table 1 [68].

Material	NDSI
Fair Skin	0.75
Darkly Pigmented Skin	0.62
Doll	0.24
Cardboard	0.14
Paper Bag	0.15
Soil	- 0.10
Red Brick	-0.01
Grass	0.53
Leaf	0.27

Table 1: NDSI Values for Various Materials [68]

Once human skin is detected and segmented from the HSI, a measurement of the melanin level is easily accomplished with the NIMI calculation shown in Equation 22 [68].

These additional clues should be used to aid the recognition of non-cooperative subjects. To demonstrate how this capability could be extended, the same simple detection algorithm is applied to an image where significant amounts of make-up have been applied to the face and the results are shown in Figure 45. The unusual return, where the makeup is applied, highlights that the individual is attempting to conceal features or change their appearance.

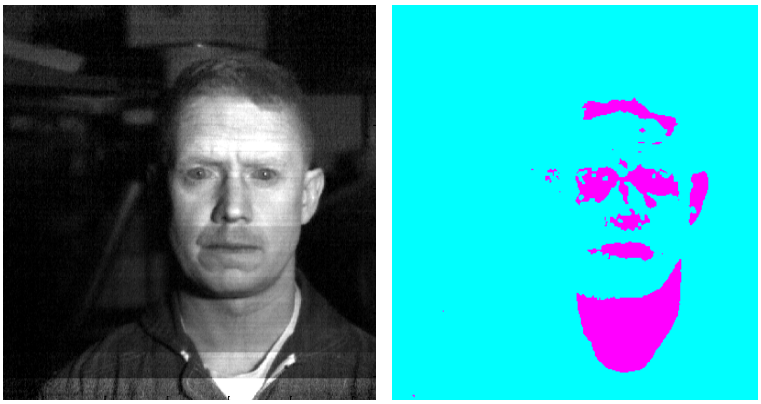

Figure 45: NIR Skin Detection When Make-up is Applied

Extending this capability further with the aid of Independent Component Analysis (ICA), the NIR HSI, shown in Figure 46, highlights the underlying veins and hemoglobin effects mentioned earlier. This capability offers complimentary value to recognition efforts and detecting spoofing attempts. Not surprisingly, this technology is starting to emerge in the leading 3-D hand geometry systems.

Figure 46: Vein Patterns in Hyperspectral Hand Image

In the medical community, this capability is being actively researched for other purposes. Paquit explored the best combination of NIR wavelengths to locate veins for the purpose of automating the surgical insertion of an intravenous catheter [69]. Using six different wavelengths, in the range of 740nm to 910nm, he tried to determine the best combination of wavelengths using linear discriminant analysis.

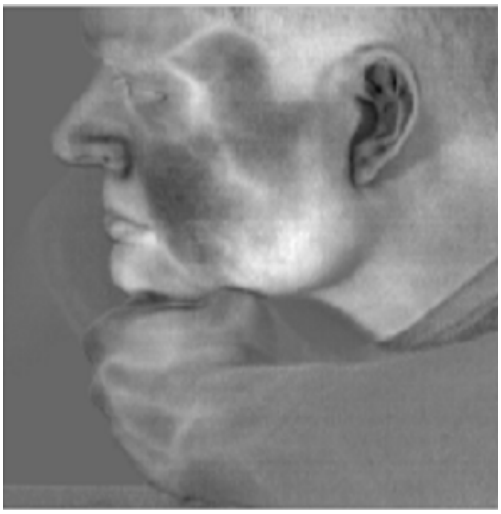

Figure 47: Vein Patterns on Hand and Face

The next illustration locates the veins on the back of the hand and possibly the face, as shown in Figure 47. Some interesting and perhaps discriminatory effects are also observed on the subject's ear lobe. Due to the unique characteristics of the ear and the stability in appearance as we age, the performance of many ear biometric applications have shown much promise as summarized by Yan [70] in Table 2.

Reference	Data Used	Dataset Size	Time Lapse	Number of G/P*	Method Applied	Earrings /Occlu.	Reported Perfor.
Chen & Bhanu [8]	3D	30 × 2	Same Day	1/1	ICP	No	93%
Hurley & Nixon [17]	2D	63 × 4	5 Month	1/1	PCA	No	99.2%
Moreno et al. [23]	2D	28 × 6	Different Days	1/1	Neural Net	Not mentioned	93 %
Yuizono et al. [31]	2D	110 × 6	Same Day	3/3	Genetic search	Not mentioned	100%
Victor & Chang [6]	2D	88 × 2	15 Month	1/1	PCA	No	73%
Choraś [10]	2D	N/A	Same Day	N/A	Feature-based	No	100%

Table 2: Ear Recognition Studies [70]

Yan [70] used a complicated process of using both 2D and 3D imagery data, conducting a color-based skin detection, followed by curvature estimation and surface curvature segmentation and finally a region classification and ear pit detection was necessary to automate the ear detection and recognition system. Chen [71] was also required to employ both modalities of color (2D) and range data (3D) to obtain competitive results for an automated ear recognition algorithm. The appearance of many of the distinctive features of the ear shown in the HSI indicates an opportunity to exploit the single modality of HSI. The complementary nature of ear and face recognition

becomes evident in most any profile view that challenges a face recognition application but often reveals a readily visible image of the ear. The fusion of the strongest features offers the best opportunity for the design of a robust identification system.

Finally, the correct question in many of these cases may not be who you are, but what you are up to. Pavlidis [72] addressed the later question in his research that used mid-IR (3-5 μm) and far-IR (8-12 μm) to sense temperature variations that accompany anxiety and fear. Many of the previously discussed methods along with this novel application can be used not just as a means to identify an individual but also as a supplemental soft biometric trait that indicates when an individual is trying to avoid detection or is exhibiting behavioral characteristics that are not easily concealed.

Hyperspectral Face Recognition

The visible portion of the electromagnetic spectrum is among the most heavily researched areas and a growing number of efforts exploring other segments of the electromagnetic spectrum have resulted in increased performance in both tracking and recognition testing. Various experimental efforts exploring the utility of IR face recognition has reinforced its utility in locating and tracking faces but the variations in thermal signatures over time and under different environmental conditions negatively impacts its performance to consistently identify individuals [90]. For this reason, many approaches fuse IR and visible images to yield the best overall performance [74].

In the NIR spectrum, Pavlidis [54] and Dowdall [91] demonstrated the ability to efficiently segment faces, a critical first step in the recognition process. This method used skin reflectance properties in NIR wavelengths that highlight human skin and aid in disguise detection. NIR wavelengths also have a larger penetration depth than visible wavelengths and are less dependent on skin temperature than thermal IR [92]. These attributes provide a more stable representation of surface features that are less susceptible to alteration.

Previous research using multispectral images identified capabilities that could be incorporated into a hierarchal framework for face recognition. One of these applications by Singh [74] looked at image and feature fusion using phase and amplitude information of several wavelengths. In this effort, Singh used Discrete Wavelet Transform to perform image level fusion of long-wave IR and visible spectrum and a Support Vector Machine (SVM) to choose either the phase or amplitude features that are extracted using a 2D log polar Gabor Wavelet. Although many of these multispectral research efforts offer valuable contributions to the field of face recognition, they do not fully exploit the span of information contained throughout contiguous wavelengths of a HSI cube.

Hyperspectral Research

Robila [93] expanded face recognition research to 120 wavelengths encompassing both the visible and NIR wavelengths, as he explored the utility of using spectral angles for comparison. He was able to distinguish between 8 test subjects using the average of 18 samples from each person's face. Comparisons were not only distinguishable between

subjects but also across different regions of the face as depicted in Figure 48. For hyperspectral signature comparisons, Spectral Angle Measurement is a common metric used in many hyperspectral software packages.

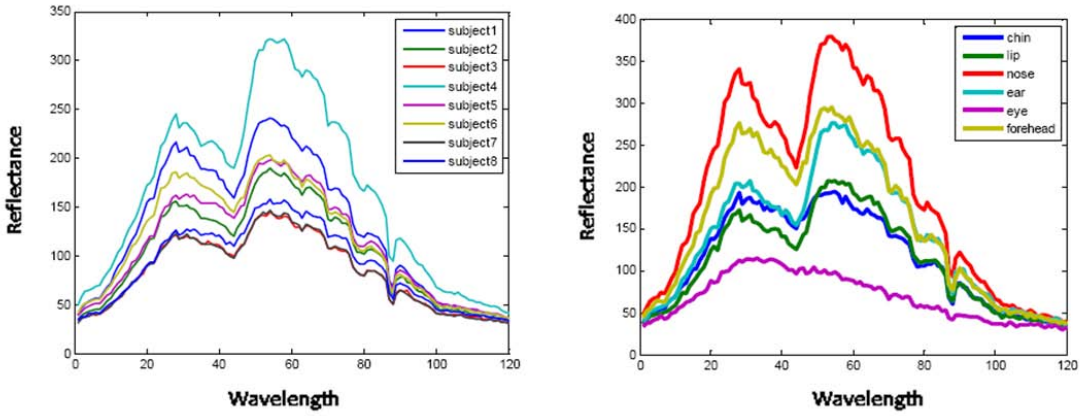

Figure 48: Spectral Signature of Various Individuals Faces and Regions [93]

The spectral angle can be calculated for two spectra, $\boldsymbol{x}$ and $\boldsymbol{y}$, in the following manner.

$$Spectral\ Angle\ (\boldsymbol{x}, \boldsymbol{y}) = \arccos\left(\frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\sqrt{(\boldsymbol{x} \cdot \boldsymbol{x})(\boldsymbol{y} \cdot \boldsymbol{y})}}\right)$$

Equation 23: Spectral Angle [93]

In Figure 49, the three segments shown (blue, red, greeen) compare the spectral angle between a test subject and the 8 individuals in the gallery from 3 different pose angles (front, 45 degree profile and 90 degree profile). When the comparitive angles are plotted for all subjects, it becomes apparent that subject number 6 has the closest spectral angle to the test subject. Robila chose this efficient technique after an earlier work that

evaluated the use of Euclidean distances, spectral correlation angle as well as spectral angle [94].

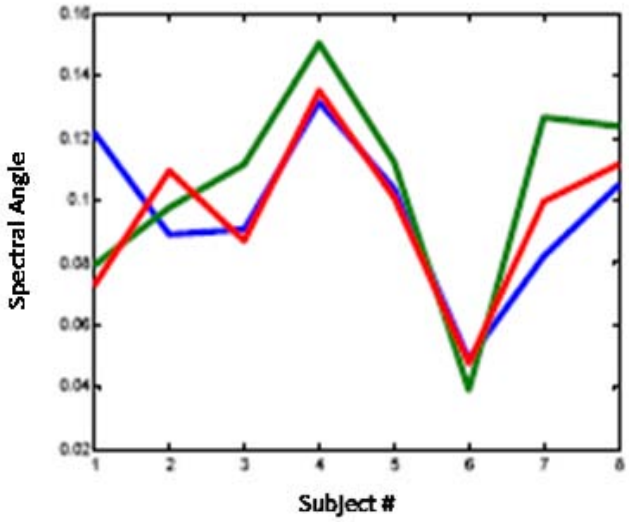

Figure 49: Spectral Angle Comparison of Eight Test Subjects [93]

Some of the attractive attributes of the spectral angle measurement is that it is always greater than or equal to zero (between zero and $\pi/2$) and invariant to scalar multiplication. The defined interval for this metric is useful in setting predefined thresholds and the metric's invariance to scalars allows two spectra under different illumination conditions to result in a spectral angle of zero. This invariance can prove useful for analyzing images under different illumination conditions that are uniform in nature.

Chou [95] used HSI and experimented with segmenting different tissue types in the human hand using Euclidean Distance (ED) and Spectral Angle Mapper (SAM). These distances are utilized in an Isodata Clustering approach that is an advanced k-nearest neighborhood method. Obtaining mixed results, an additional technique was

employed to separate fingernail regions from the rest of the image. A sample variance was computed over a small spatial neighborhood using a Round Filter. The imagery used in these initial experiments includes both visible and NIR spectral ranges.

Pan has utilized hyperspectral spectral signatures of skin to accurately identify 200 individuals in several related efforts [96], [92], [58], [75]. His methodology capitalized on the spectral signatures obtained from different tissue types to create a robust recognition capability when faced with pose, expression, and illumination variations. Pan's methods utilized an initial identification approach that was based solely on the Mahalanobis distance between spectral signatures manually extracted from each subject's hair, forehead, cheeks, lips and chin. Performance results for various demographics, poses, and expressions were identified in this research.

In a subsequent effort, Pan [75] explored the benefit of using a holistic similarity distance measurement using eigenfaces at each wavelength. The Mahalanobis Cosine distance was measured at each wavelength and the results summed for an overall metric of similarity. Additional methods including using PCA to sort the bands and compare the cumulative Mahalanobis Cosine distances as additional bands were added. Finally, a spectral face representation was created by assembling a face image by incrementally selecting each pixel from the following band and then repeating until the entire 2D image is created. This method attempted to preserve both the spatial and spectral properties of the hyperspectral face cube. All methods achieved very successful results with cumulative match scores greater than 0.91 for rank 1, which simply means that 91% of the first matches made by the algorithm were correct.

Since the work of Pan, one additional research effort in HSI has surfaced. Elbakary [97] used the K-means clustering algorithm to segment the skin surface in HSI and then match like signatures between test subjects using Mahalanobis distance measurements. This effort was useful in demonstrating the potential to automate the selection of skin segments and the associated spectral signature for matching.

Reference	Author	Approach	Image	Database	Contribution
[96], [92], [58], [75]	Pan (2003-5)	Distance comparison using Mahalanobis distance of spectral signatures; PCA and Multiband; image projection into desired illumination space	Hyperspectral [700-1000nm] 31 bands 468x494	University of California at Irvine 200 subjects	Demonstrated robustness to variations in illumination, expression and pose
[97]	Elbakary (2007)	K-means clustering and Mahalanobis distance comparison	Hyperspectral [450-1100nm] 65 bands 640x480	CMU Hyperspectral 48 Subjects	Automatically clustering of face tissue by spectral angle
[93]	Robila (2008)	Spectral Screening for the identification of tissue types and individuals	Hyperspectral [400-900nm] 640x640	Montclair State University 8 subjects	Ability to distinguish individuals and tissue types based on spectral angle

Table 3: Summary of Hyperspectral Face Recognition Research Efforts

Despite the vast amount of face recognition research already accomplished and the growing popularity of HSI for remote sensing applications, there have been only a few efforts combining the two technologies. Table 3 summarizes the hyperspectral face recognition research accomplished to date. These efforts have given us the foundation to exploit HSI for face recognition. From these findings, an automated and robust

application can be constructed in a relational hierarchy that that combines proven capabilities from the spatial, spectral, and temporal domains, not so different from the manner in which human recognition is accomplished.

Fusion

Even with the distinctiveness that accompanies every human being, no single metric or feature has demonstrated the ability to identify all individuals in both controlled and uncontrolled environments, and certainly over large populations. Therefore, it may be due to the infancy of biometric systems or the sensitivity that people have about their privacy rights, that many commercially employed systems are typically limited to a unimodal application. Regardless of the biometric chosen, the single mode will contend with the challenges of noise in the sensed data, intra-class variations, and inter-class similarities [16]. Multi-biometric systems can address some of the challenges facing unimodal biometric identification but they come with added computational demands, complexity, and costs [73]. Therefore, the design and implementation of a multi-modal solution should carefully weigh the cost-benefit of these trade-offs.

Solutions that incorporate several modalities and multiple sources of information are able to increase the population coverage and counter spoofing attempts, all with a proven increase in performance [16]. However, with these multimodal methods comes the requirement of close proximity to the subject as well as their compliance. Unfortunately, the difficulty with identifying a subject is occasionally due to a lack of full

cooperation and so the expectation of voluntary participation and immediacy could be wishful thinking.

An alternative to this challenge may be to fuse contextual information in an architecture that enhances effectiveness. Regardless of the eventual solution, there remains the task of determining the best combination of diverse and complimentary information from different domains. In order to discover the best fusion strategy, a review of common alternatives is necessary. The basic levels at which biometric fusion occur are at the data level, feature extraction level, matching score level and the decision level. For the following discussions, fusion is defined in the following manner. If, φ is a measure of performance, then fusion using the rule *r* is:

$$\varphi(r(I_1, I_2)) \geq \max\{\varphi(I_1), \varphi(I_2)\},$$

so that,

$I_3 = r(I_1, I_2)$ is the fused information.

Fusion Strategies

At the beginning of the process, data fusion can occur as images taken from different wavelengths or sensors are combined using image and signal processing techniques [104]. A common example seen in the literature is combining the tracking advantages of IR sensor data with the spatial detail of visible imagery to achieve a level of performance higher than either of the systems used in isolation. Illustrating the advancement in sensor technology, Varshney [104] was able to combine images from an

IR and millimeter wave (MMW) sensors to obtain an image that was able to more accurately depict a concealed weapon (see Figure 50).

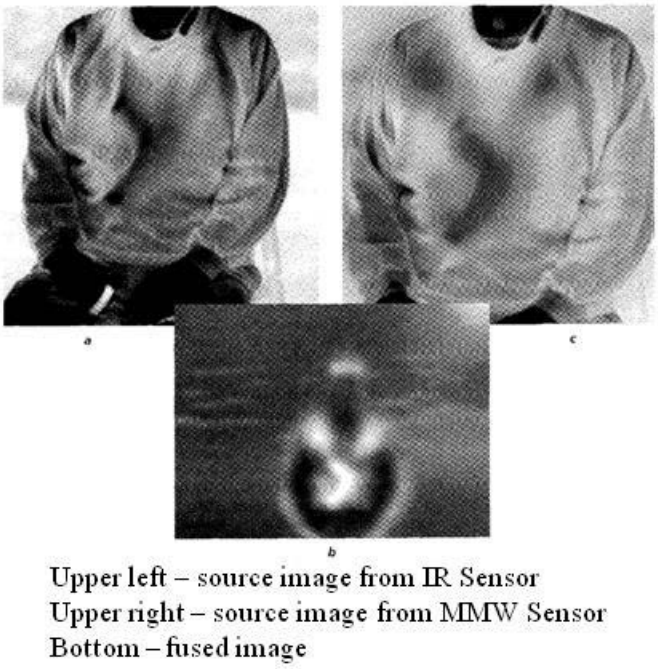

Figure 50: Data Fusion Using IR and MMW Images [104]

At the feature extraction level, various features are collected from different sensors and combined for a higher dimensional feature vector. A cautionary note by Jain [73] notes that not all biometric features are easily combined. For example, the eigenface weightings from a face recognition system and the minutiae traits from a fingerprint are not intuitively similar and therefore difficult to merge [73]. When these features are able to be combined, they are often characterized by increased dimensionality that makes feature reduction techniques necessary in this larger hyperspace. An example of a successful fusion effort at the feature level was the work of Marcel [46], who effectively combined a grayscale image and a RGB distribution vector as a skin feature.

At the decision level, fusion can be accomplished in a straightforward manner as results (accept or reject) are tallied with a simple voting algorithm. This level of integration can be seen as limiting as a preponderance of the information has already been discarded and is no longer available for consideration. The drawbacks of feature and decision level fusion have made the relative ease of matching score fusion a popular approach [73].

Fusion at the matching level is most often seen with resulting scores being normalized before combining them in simple averaging or ranking techniques. To illustrate how easily one of these techniques can be applied in hyperspectral face recognition a recent research example is used. Pan implemented a simple sum rule to combine scores, in this case Mahalanobis Cosine distances, $D_{\mathbf{u},\mathbf{v}}(w)$, at each wavelength ($w$) as shown in Equation 24 [36], [75].

$$D_{\mathbf{u},\mathbf{v}}(w) = \frac{-\mathbf{m}_w \square \mathbf{n}_w}{\left|\mathbf{m}_w\right|\left|\mathbf{n}_w\right|}$$

Equation 24: Mahalanobis Cosine Distance [36][75]

The two images (**u**, **v**) are used to obtain a distance measurement at each wave length ($D_{\mathbf{u},\mathbf{v}}$) and then combined for an overall evaluation of face similarity ($\mathbf{D}_{\mathrm{u,v}}$) shown below in Equation 25 [75].

$$\mathbf{D}_{\mathbf{u},\mathbf{v}} = \sqrt{\sum_{w=1}^{W} (1 + D_{\mathbf{u},\mathbf{v}}(w))^2}$$

Equation 25: Sum of Mahalanobis Cosine Distances [75]

In addition to the options of processing levels at which fusion can be implemented, there is a variety of architectures that can be selected. Biometric systems can be combined in an ensemble approach where multiple classifiers act on the same problem in a redundant manner [76]. An alternative to the ensemble approach is the modular approach, where the overall classification problem is broken into smaller subtasks with the final classification dependent on inputs from all modules used [76]. Both ensemble and modular approaches can be either combined in a serial sequence or applied in parallel. There are many options available when implementing fusion in a multimodal biometric system.

Fusion approaches can range from the simple to the complex, but the intricacy of these algorithms is not necessarily an indicator of their effectiveness. At the feature level, a simple if-then-else approach can be applied in a serial sequence to evaluate and filter possible identities feature by feature. This type of approach has been proposed in the design of a security screening application where watch list candidates are processed and separated from the general public using queuing theory fundamentals to enhance the efficiency of screening passengers [77].

At the matching level there are many options for combining the numerical scores. Some of the most intuitive approaches are variations of summing and averaging rules. Indovina [78] found in his attempt to combine face and fingerprint scores that a Min-Max normalization followed by a simple sum rule out performed commercial off the shelf unimodal biometric applications. Ross [79] likewise found in his research, combining face images, fingerprints and hand geometry, that the basic summing rule was the most effective among the strategies tested. Averaging rules, as the name implies, combine all

scores as a representative mean, but can be tailored to more sophisticated approaches that weigh each individual score based on the relative importance of each attribute.

Weightings can be determined by the accuracy of the classifier used, by the quality of the image, or by predetermined values that minimize overall classification error. These weights can be obtained through classifier training or adaptively adjusted to improve performance. Indovina's [78] review of weighting and fusion schemes mentioned earlier, found that for closed populations the best approach changes based on the ability to re-sample the same individuals. His Quadric-Line-Quadric adaptive normalization and user weighting strategies outperformed the sum fusion method as repeated sampling and statistics were employed for increased performance.

The use of a ranking rule is another popular technique for combining matching scores. The Borda Count Method is a common ranking rule where each score is assigned a ranking (highest rank to the best match), and then these ranks are summed with the final classification going to the overall highest ranking. Figure 51 depicts this approach as the probe subject is matched to template 2 due to its similarity distance score in both the PCA and LDA application. In this example, template 2 achieves the second highest matching score in the PCA system and top matching score in the LDA system making it the overall selection for a face match. It should be noted, the conflicting decision that a summing or averaging approach would present in this case as template 1 actually offers a lower average score.

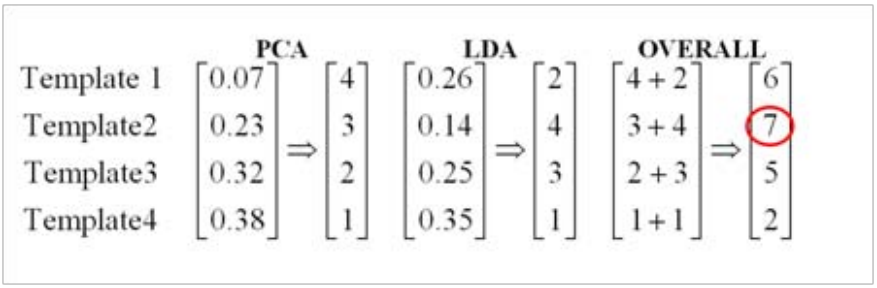

Figure 51: Borda Count Method

When fusion is applied at the decision level, much of the data processing has already been accomplished, so the remaining task can be very straightforward. The decision outputs from the biometric systems are routinely combined in a voting rule. A majority-voting rule selects the most frequent class label that appears for a probe. The computational simplicity of these elementary fusion techniques can make them an attractive choice.

Theoretical Framework

To formulate the combined classification system proposed for a contextual face recognition system, a derivation from Jain [80] can be used to determine the probability of identifying a test subject (ω_i) given both a primary biometric feature vector (x) and a secondary biometric feature vector (y). In this research, two separate systems produce independent feature vectors, and with the help of Bayes' rule, the following calculation for a final matching probability is derived (Equation 26) [80].

$$P(\omega_i \mid x, y) = \frac{p(y \mid \omega_i)P(\omega_i \mid x)}{\sum_{i=1}^{n} p(y \mid \omega_i)P(\omega_i \mid x)}$$

Equation 26: Probability of Fusing Biometric Features [80]

Although, this approach may helpful from an application standpoint, it can also be constructive in determining the error bounds for proposed fusion rules. Previous work accomplished by Schubert [81] used Boolean label-fusion rules to determine the combined ROC curves for these systems. Wang [82] was able to derive and predict the expected performance of simple 2-set combinations of biometric classifiers for various score fusion rules using the area under the ROC curve, likelihood ratio and a discriminability metric [82].

If a majority voting rule is chosen, certain assumptions must be met to ensure a successful design and implementation. Assuming the independence of errors among the various classifiers and an equal error probability of $p < 0.5$, the error of the majority-voting rule would be monotonically decreasing [83]. If the correct classification is made by a system with likelihood (1- p), the chances of seeing exactly k errors among N classifiers is calculated using Equation 27.

$$\binom{N}{k} p^k (1-p)^{N-k}$$

Equation 27: Probability of *k* Errors for Majority Fusion Rule

The likelihood of the majority rule error shown in Equation 28.

$$\lim_{n\to\infty}\sum_{k>N/2}^{N}\binom{N}{k}\,p^{k}(1-p)^{N-k}=0$$

Equation 28: Likelihood of the Majority Rule Error

By induction, for either odd or even *N*, if each classifier gets the right answer more than half the time, assuming independence, then as more classifiers are used, the likelihood of an error decreases when using majority rule. The respective error cost for the optimum Bayesian fusion of *N* sensors, or from another perspective *N* classifiers, is derived in [84], [85] and is show in Equation 29.

$$\sum_{i=1}^{N}\left[u_i\log\left(\frac{1-F_{RR_i}}{F_{AR_i}}\right)+(1-u_i)\log\left(\frac{1-F_{RR_i}}{F_{AR_i}}\right)\right]_{u_g=0}^{u_g=1}\ \text{¤}\ \log\left(\frac{C_{FA}}{2-C_{FA}}\right)\text{c}$$

Equation 29: Error Cost for the Optimum Bayesian Fusion of N Sensors [84][85]

The false rejection rate (F_{RR}) and false acceptance rate (F_{AR}) as well as the associated cost (C_{FA}) are incorporated into the derivation along with the local sensor decision (u_i) and global decision (u_g). This result is beneficial given the 2^{2^N} possible fusion rules.

Value of Diverse Classifiers

When combining classifiers it is reasonable to assume that there should some diversity among them if their combined results are to offer some advantage. This advantage can obtained through the accumulation of independent information with different types of error or noise canceling out one another. Likewise when it comes to

classifiers, Kuncheva [86] discussed the importance of diversity among an ensemble of classifiers if the collection is to be successful. This cumulative knowledge is sometimes cited as an intuitive advantage in the human decision making of crowds [87].

Somewhat surprisingly, the higher performing classifiers may be the ones discarded in lieu of more diverse classifiers for better overall performance of an ensemble such as AdaBoost [86]. With this in mind, the evaluation and inclusion of diverse classifiers can be vital to the overall performance of the ensemble. The correlation of classifier outputs or the evaluation of their entropy can be helpful in determining the right combination of classifiers to include in a fusion ensemble [86].

In the course of discussing fusion strategies and entropy, the key concept is using the mutual information of classifiers in a way that reduces the uncertainty. In an effort to quickly review the basic tenets of information theory, entropy is a measure of a random variable's uncertainty [88]. Mutual information on the other hand is the amount of information that is shared between variables [88]. The reason these two concepts are of such interest is because of their relationship and the impending result of maximizing the mutual information which is to minimize the probability of classification error.

The critical step will be to apply these considerations to HSI in a beneficial fusion strategy. Looking at recent findings from face recognition research starts to give us an indication on how this may be done. Bowyer [89] explored several fusion strategies and modality combinations in his efforts at combining grayscale images, IR, and three-dimensional shape imagery.

Bowyer confirmed that increased performance is achieved when different types of imagery (multimodal) are combined. He also investigated whether multiple images of

the same type, in this case grayscale, would provide added performance [89]. By implementing a weighted score fusion strategy for the multi-sample approach, he was able to achieve comparable results to that of a multimodal system. Ultimately, Bowyer surmised that both multi-sample and multimodal approaches would reach a performance limit and that in order to achieve higher performance these approaches would have to be integrated [89]. This research will attempt to build upon this concept by combining the various spectral wavelengths and multiple images for maximum confidence in identifying a subject.

IV. Research

In order to develop and test various methodologies and hierarchies, a hyperspectral face database needed to be created or acquired. A few initial image cubes were able to be built with the use of local equipment. These test images were collected using two separate cameras, SOC 700 and 720, spanning the frequency ranges of 400-900nm and 900-1700nm respectively. The continued use of this equipment to develop a full database presented several challenges.

The first of these obstacles was the ability to combine these images and the associated complexity of image registration. The second was obtaining an available and functional laboratory with the appropriate lighting equipment. Finally, there was the anticipated obstacle of receiving research permissions for conducting this study using personal images of people's faces with the related privacy concerns.

Although not part of this research effort, the privacy concerns of biometric systems are very real and should be a significant area of study before fielding any operational system. The reluctance of the general populace to allow these systems to be integrated into public life is one of the most important considerations facing the implementation of any system. This sensitivity had a significant impact on the decision to not release a previously developed database for government research at any classification level, regardless of additional disclosure restrictions and agreements.

Image Data

For this research effort, there are several hyperspectral databases cited in the literature that could be used for testing. The first dataset was collected at the University of California at Irvine (UCI). The UCI Database contains NIR images from 700nm to 1000nm containing 31 spectral bands and a spatial resolution of 468x494 pixels. This data contains images from 200 subjects comprised of a diverse population of gender, age, and ethnicity. Each individual has seven images to include two front view (with neutral expression), another front view with expression, and four images from side orientations (45 and 90 degrees from left and right sides). A sample of this data is taken from one of those publications is shown in Figure 52.

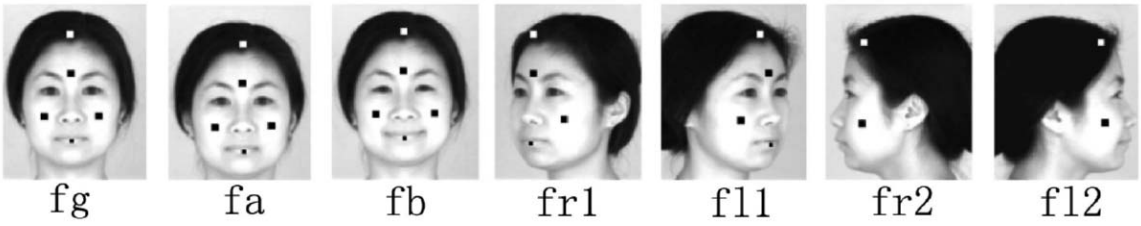

Figure 52: UCI Database – Face Image with Expressions and Rotation [96]

Unfortunately, despite the numerous publications that were based on this government funded project, the UCI authors could not locate the underlying data that this research was based on. Attempting to acquire the data at the source, it was found that this work was sponsored by the US Defense Advanced Research Projects Agency (DARPA) under the Human Identification Program (HID). U.S. Government program funding was provided both through DARPA's HID Program and through AFOSR Grant

F49620-01-1- 0058 and NIH Grant RR01192 [96]. Despite the funding source and the research area supported, DARPA was unwilling to release the data under any classification or disclosure restrictions. Multiple attempts to contact the current project manager and supporting contractor were unproductive. The sensitivity of government agencies to acknowledge and facilitate research in the field of biometrics would be a reoccurring theme throughout this research effort.

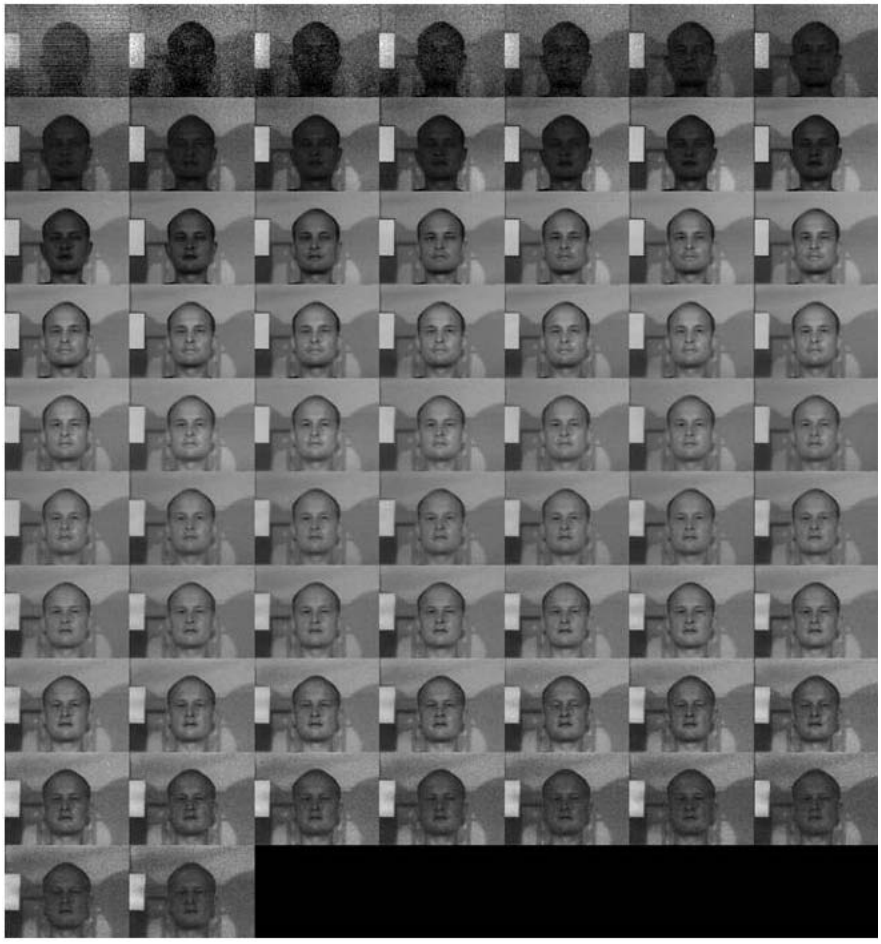

Figure 53: CMU Hyperspectral Database

The second dataset for consideration was collected by Carnegie Mellon University (CMU) [98] and was graciously provided by the well known computer vision researcher Dr. Takeo Kanade. The CMU Database contains Visible and NIR images from 450nm to 1100nm containing 65 spectral bands and a spatial resolution of 640x480 pixels (Figure 53). The original dataset contained images from 54 subjects comprised of a 4 front images with illumination from 45 degrees left, center, 45 degrees right, and an image with combined illumination. The data acquired with the cooperation of the CMU staff was a subset of the original data with as many as 48 subjects.

Another dataset that is publically available for research is a multispectral database made available by the Equinox Corporation [99]. Since this data is not a hyperspectral dataset, the ability to sample, test, and incorporate the value of spectral signatures would be lost. This data, also funded through the previously mentioned Human at a Distance program, is available online, and is depicted in Figure 54. These separate images cover the following wavelengths, short-wave IR (SWIR, 900-1700nm), medium-wave IR (MWIR, 3000-5000nm), and long-wave IR (LWIR, 8000-12000nm).

Figure 54: Equinox Multispectral Face Database [99]

A multispectral database from the University of Oulu in Finland was also obtained as another research option [100]. Designed primarily for the study of color image analysis, it contains color face images from 125 people under 16 different camera calibrations and types of illumination. Three skin reflectance spectral signatures (400-700nm), measured from the cheek and the forehead, are provided for each person. This data is illustrated in Figure 55.

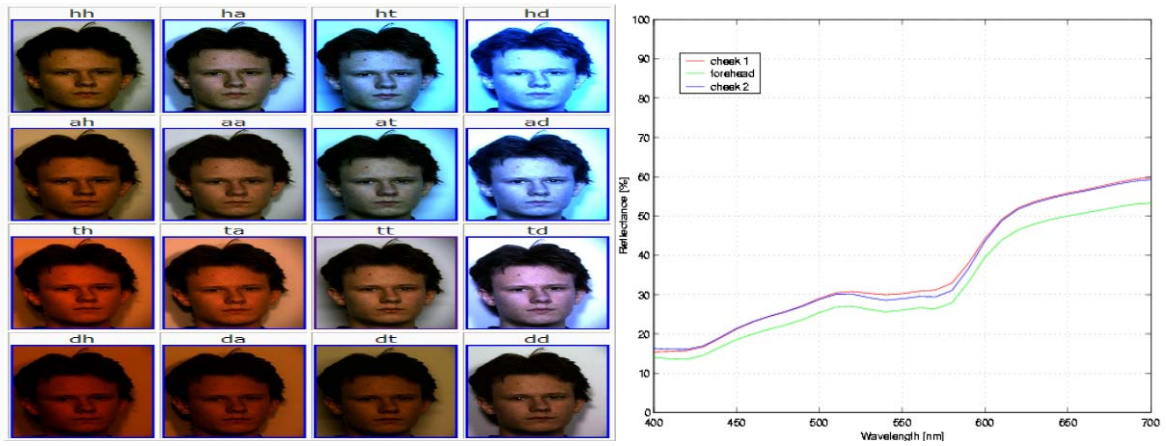

Figure 55: Oulu Color Image Database [100]

At the end of this research effort, additional upgrades in sensor equipment have provided an impressive capability to capture HSI from 400nm-2500nm with a single camera, removing the additional image registration issues. The new hyperspectral video recording equipment also provides a real capability to explore temporal dimension and value of video images. This rapid pace of evolving equipment capability indicates the eventual progression and direction of this maturing technology.

QUEST Methodology

Qualia can be defined as a representation of the physical environment or a facet included in ones internal representation of the world around them. It is the part of your internal representation that is introspectively unavailable [101]. Rogers [102] refines this definition by adding that it is any discernible aspect of the illusory Cartesian theater. It is argued that by combining different qualia into a meta-representation, sensory inputs can be integrated into a world model that is adaptable and efficiently functional.

The Qualia Exploitation of Sensor Technology (QUEST) methodology attempts to develop a general-purpose computational intelligence system that captures the advantages of qualia-like representations [103]. A guiding principle of QUEST highlights the use of qualia that map sensory input to more useful and efficient states. The functional requirement for a QUEST system is that it possess the ability to detect, distinguish and characterize entities in the environment [103]. In order to build a QUEST system for our task it is important to develop and understand the concept of an agent.

An agent, depicted in Figure 56, takes a subset of stimuli from the environment and processes this into information. An agent has knowledge of other agents and transmits selected aspects of its information representation to other agents. Agents transmit stimuli upward in higher levels of abstraction and transmit some information downward providing context that can influence lower level agents. An entity uses a set of agents to create an internal representation of its environment. The internal representation is formed through the collective knowledge of the agents.

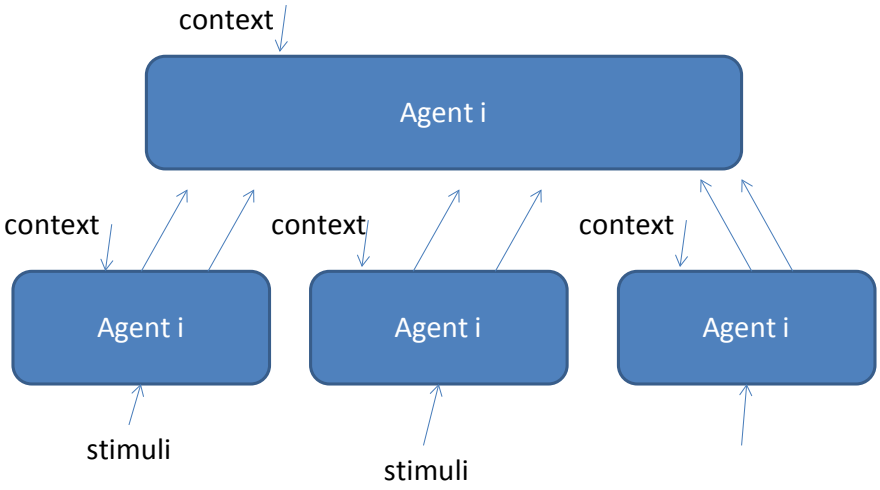

Figure 56: Agents and Level of Abstraction [102]

In this research, the relevant attributes are compromised of biometric characteristics and contextual information across the electromagnetic spectrum. Rogers [103] states that the concept of context is the subjective representation of an entity as it exists in the world, an abstract characterization or general mental concept that helps us identify an object. The combination of detailed fiducial features from the stimulus space and higher-level abstracted biometric features create this context. As Rogers [103] reminds us, memory is not just the storage of prior experiences but an iterative and constructive process that is utilized in perception, recollection, and imagination. In the human recognition system, the mind stores data not so much as sensory numbers but as relative comparisons to prior experiences, that can change over time. For the purposes of developing a face recognition system, the relative comparisons should serve an important role in refining the search space and guiding the search process.

The connections or links in our system will provide the logic and determine the effectiveness of context. There are many choices available to choose from and links can

be the connection of internal and external facial features that have proved so important in human recognition research. The links chosen can help incorporate higher levels of abstraction such as important soft biometric cues. The links can be the connection between spatial and spectral information. Spectral information can be used to segment facial features such as hair, skin, and eyes and then this information can be transmitted to spatial recognition agent or conversely a spatial segmentation approach can supply knowledge useful for spectral matching.

Selected biometric features and the links between them that create the contextual backdrop should be continually refined in a reoccurring process that reduces the uncertainty in authenticating or identifying a person. By developing a rational feedback loop that iteratively looks at the problem an adaptive behavior can emerge. The process must be efficient to explore the vast amount of sensory data contained in HSI, filter relevant aspects to create information and then combine and update this information to create knowledge.

Soft biometrics in this research are not merely labels but instead, vectors for a process that generates information and reduces the search space. Higher-level abstraction is a benefit that can help guide classical recognition approaches. Without this assistance, a recognition system could be characterized as autistic without an ability to connect important semantic categories. In order to create a theory of mind, the developed methodology may be well served by mimicking the human recognition system that incorporates an overt process that enables humans to identify faces and the covert process that allows individuals to connect semantic information for confirmation [10], [11].

The intelligent use of biometric cues can enable the use of spatial, spectral, and temporal data much the same way the human eye and brain operate interactively to utilize stimuli. We intend to build a hierarchy to enable the fusion of independent systems to create an efficient representation of a hyperspectral face image for purposes of identification. One of the goals of this research is to integrate the concept of qualia to efficiently exploit sensory data in a method that transitions from general to specific characteristics in a hierarchal architecture. The Qualia Exploitation of Sensor Technology (QUEST) methodology applied in this research will attempt to replicate contextual and temporal information in an adaptive feedback loop. The face recognition problem offers an attractive test bed to investigate and develop evolving QUEST concepts that can promote both the field of biometrics and generic ability to detect, distinguish, and characterize entities in the environment.

Illustrative Example

One possible application of a face recognition system is to locate and identify a person of interest from a national watch list using a collection of face imagery data, hyperspectral in our case. This can be aided by an associated soft biometric description (i.e. male, black hair, light complexion). A secondary requirement of our system is to monitor individuals trying to circumvent detection.

Figure 57: Skin Component for Face Detection (Theoretical) [106]

In this setting, our first task is to detect human skin and subsequently faces in HSI (Figure 57). Using the Normalized Differential Skin Index by Nunez [54], or similar approach used by Pavlidis [54], face images can be quickly located and extracted. Although the spectral signature of human skin is distinctive from other materials and unique among population groups, it may only be useful to narrow our candidate pool, but not adequate for a confirmed identification.

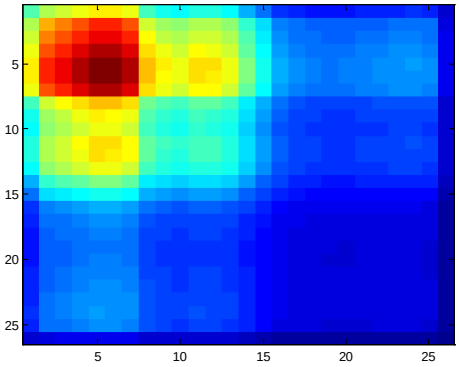

Figure 58: Hyperspectral Data Cube Covariance Matrix

Using captured face images, an initial investigation of the underlying data and covariance matrix (Figure 58) indicates that a large portion of the variance is contained in several of the first several principal components. Discarding the components associated

with lighting variance and retaining the remaining principal components and reconstructing the image would reveal a stable face representation suitable for a holistic matching approach such as eigenface. Figure 59 shows an example of an image portrayed by this reduced data.

Figure 59: First Principal Component of HSI

Understanding of the negative impact that hairstyle changes can have on recognition performance, hair segments will be removed in order to process only the face portion of the image. Matching scores based on an average hair signature could be fused into the recognition process later. Based on the initial results of the holistic matching, the face space will be tuned to more effectively highlight the unique characteristics possessed by our person of interest. Using insights from Johnston [8] and Valentine [7], a similarity measure will define a specified face space where comparable individuals can be closely evaluated or alternatively a projected space that places faces of interest in a sparse region where delineation among neighboring faces is straightforward.

Other principal components or spectral signatures, illustrated in Figure 60, could be used to segment the eyes and lips in addition to the hair. Subfeature approaches such as spectral matching, eigeneyes and eigenmouth can be used to match these features to a list of candidates. Appropriate weighting values will be assigned to these subfeature scores based on their relative value to recognition performance. Using findings from the study of human recognition, we can expect eye features to be associated with higher weights while the mouth would obtain a lower initial weight. The matching score of our subfeature method will have to be fused with holistic scores for maximum effectiveness.

Additionally, soft biometric traits can guide the candidate filtering to fit the description of our person of interest. The unique spectral properties of human skin and hair can be used in conjunction with a spectral angle metric to filter out large parts of the gallery. Additionally, DFFS can be used to selectively choose similar type images (among the gallery) based on pose, illumination or some other selected linear subspace, similar to the view-based methods discussed earlier from Pentland [24].

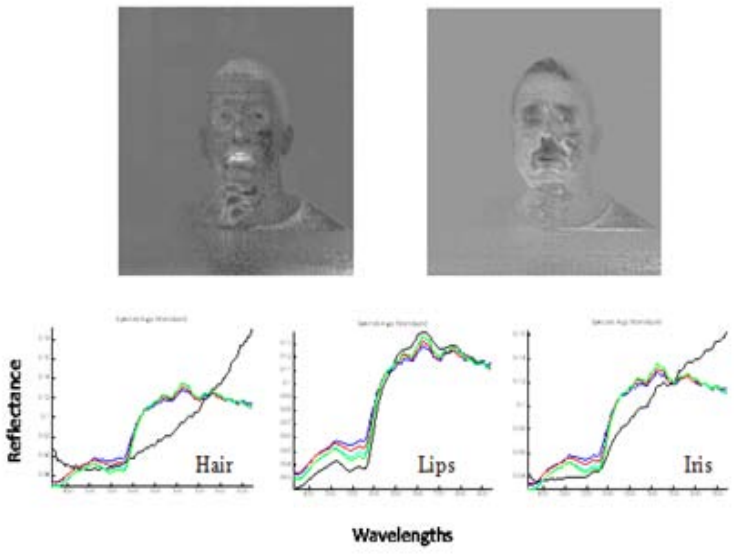

Figure 60: Additional Principal Components and Spectral Signatures

In this illustrative example of contextual hyperspectral face recognition sensory information can be collected, processed, and utilized in a manner that aids classifier performance and the overall recognition capability. The overall process incorporates a feedback loop that allows progress to continue or alter direction when there is none. The process overview is depicted in Figure 61.

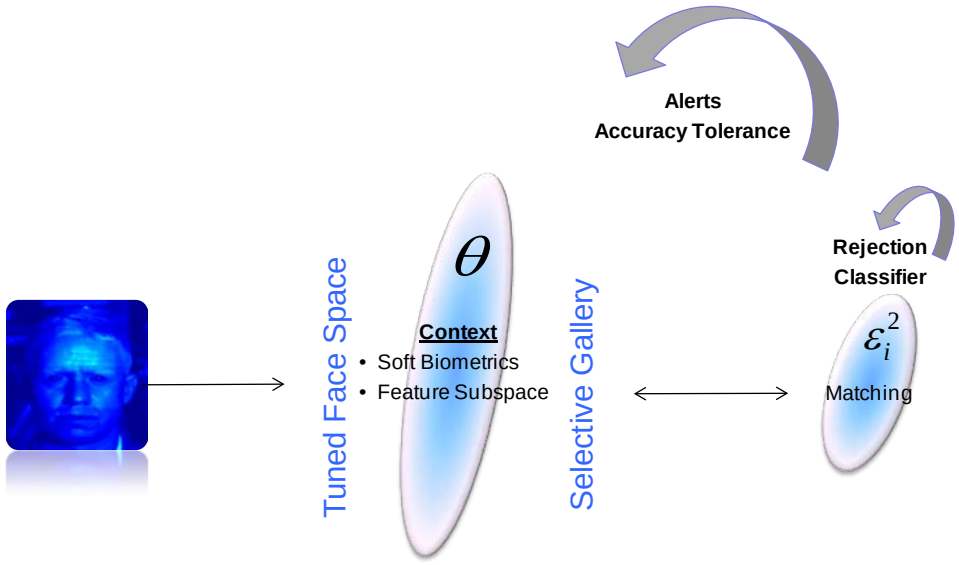

Figure 61: Initial Hyperspectral Face Recognition Process

Without the insight that experimental results will eventually provide, the initial framework is set up to allow the exploration of opportunities for QUEST concepts for hyperspectral face recognition. This general structure allows contextual benefits to influence both the comparison face space and the management of gallery candidates for matching. With this overarching approach, many important questions remain unanswered. How is the comparison face space developed and adjusted? Is the entire gallery of faces used initially? As the gallery is culled, how is the subset of candidates

chosen and adjusted? How are the scores managed, by distribution and range? Soft descriptions can theoretically guide these decisions but the construction of an application requires that these questions have an answer.

Published research from the cognitive sciences helps us to assemble a common understanding of the highly complex process of human recognition. One of the first observations is that face processing is accomplished with such speed that the initial part of this task is most probably accomplished through a predictive and direct neural process [107]. Low spatial frequency information is extracted and used for initial object recognition in a top-down process and high frequency information is added through bottom-up feedback streams [108]. Additional memory associations are used to help derive future expectations or predictions [108]. This process reoccurs with additional information and expectation updates until a level of distinctiveness and perception is achieved that allows a face to be identified.

There is opportunity to connect this cognitive research with the computation design using common transformations in both the spatial and frequency domains. Many of these transformations to include: PCA, ICA, the discrete cosine transform and the discrete Fourier transform have been used in various face recognition approaches but not with HSI. By controlling the initially retained information with context and managing the process by which supplemental information is added, the concepts described in the neuroscience literature can be evaluated experimentally.

Research Goals

The purpose of this research can be summarized by the following four goals:

1) Extend the existing research in the area of Hyperspectral Face Recognition
2) Define and explore a framework for contextual based face recognition
3) Devise a method of generating subjective representations to evoke perception or qualia-like solutions
4) Through this methodology improve the performance, efficiency and robustness of classical face recognition methodologies

Dr. Zhihong Pan has set the foundation for the study of this area with his doctoral work at UC Irvine and his publications that followed [96], [92], [75], [58]. Within the framework for contextual based face recognition, it is hoped that common weakness of face recognition (uniqueness, performance, and circumvention) as well as the reoccurring challenges of variation in illumination, pose, and expression are mitigated through this architecture. Finally, the most forward reaching goal is to provide a foothold as a driver problem for a general-purpose machine recognition ability described in the QUEST Tenets [103].

Data Exploration

Among the available data options, only the CMU database provided a true hyperspectral database that was comprehensive enough to develop and explore our intended areas of research. The initial exploration of the CMU database confirmed the difficulties with this early attempt to capture face images hyperspectrally. Denes [98] noted in his research report that the prototype spectropolarimetric camera used was

subject to stray light leaks and optical imperfections. In summary, he said, “better face recognition clearly requires higher definition through a more sensitive, low noise camera or through higher levels of illumination.” [98] This experimental data would provide an ideal challenge for our algorithms and fusion strategy development.

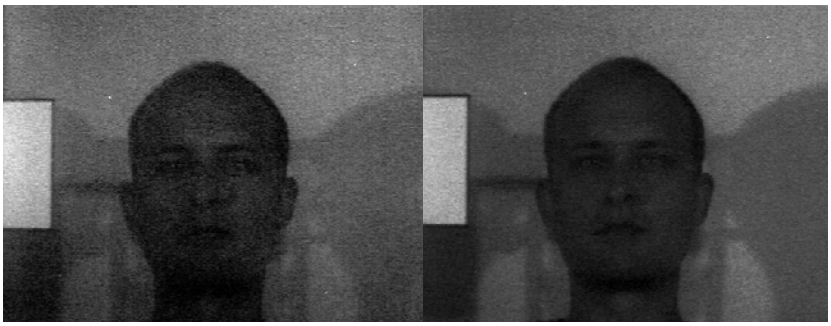

Figure 62: Blue and Green Wavelengths of an Image

The images proved to be every bit as noisy as promised, particularly at wavelengths below 600nm containing the important blue and green bands of the visible electromagnetic spectrum. To illustrate this point, sample images in the blue and green wavelengths are shown in Figure 62. The variability of this noise is also prevalent in the spectral signatures and is illustrated in Figure 63, as skin signatures for a small sample from a subject’s forehead is plotted. The spectral signatures from the 17x17-pixel sample from a relatively uniform portion of the image foreshadow the unpredictability expected throughout the database. The calculated variability is shown on the left side of the figure and unfortunately, the greatest variability happens to be present in the visible wavelengths of the image. This image characteristic should have a negative impact on any spatial recognition algorithm used on this data.

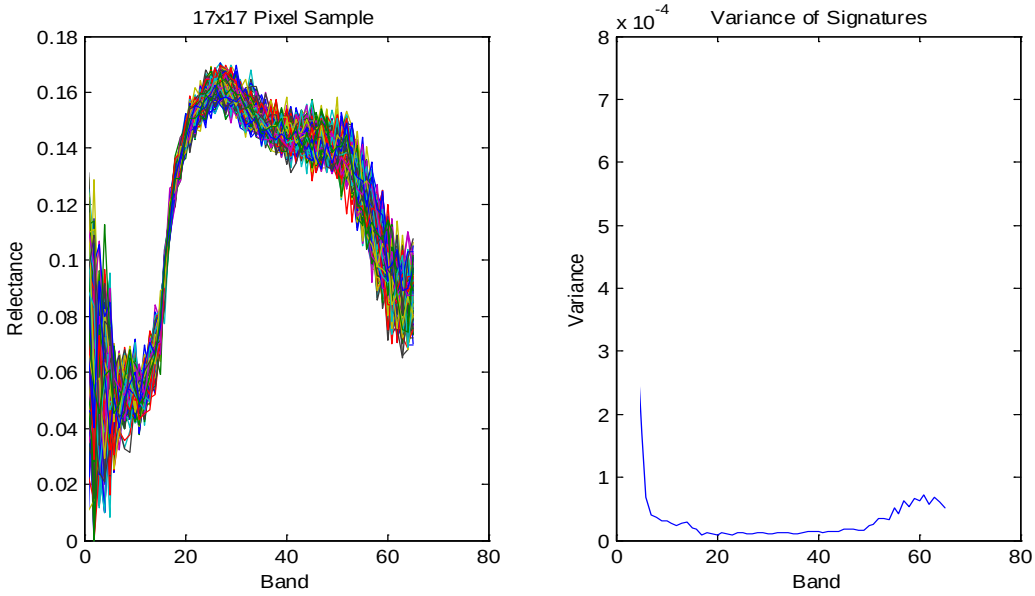

Figure 63: Variability of 17x17 Pixel Spectral Sample from Forehead

In addition to these instrumentation challenges, there were some additional complications introduced by the test subjects during data collection. Two of the more common examples are depicted below in Figure 64 and Figure 65. In the first example the test subject's head, face, mouth and eye movement occurred during the image capture as the camera sequentially cycled through the 65 wavelengths. In the second example, the test subject changed appearance by donning sunglasses during the second image session. Eyeglasses are donned by three subjects during the second session that is used for our test images. The movement and susequent changes in appearances and hair styles sometimes resulted in a partial face images similar to those expected in an uncontolled environment. No efforts to make registration corrections for the sequential images will be made and will have a detrimental impact to the performance of the identification system.

Figure 64: Movement of Subject during Image Capture

Figure 65: Appearance Changes between Training and Test Sets

Preprocessing

The importance of preprocessing images before testing is significant and the amount of effort both manually and computationally is usually not discussed at much depth. Many applications use images accompanied by the manually selected coordinates of the eyes that is used for alignment and sizing purposes. This upfront effort can be time consuming and assumes the involvement of human recognition at the onset of the process.

One such application is the Colorado State University (CSU) Face Recognition Evaluation System [36]. This application provides the capability to use a standard set of face recognition algorithms including the PCA based approach Eigenfaces, a combined approach using PCA and LDA, a Bayesian Intrapersonal/Extrapersonal Classifier, and a computationally intensive Elastic Bunch Graph Matching algorithm [36]. This testing environment also provides statistical methods for comparing the performance of these various algorithms. These algorithms, however useful, can take a long time to run in this testing environment with as much as a day for the first 3 algorithms and 5 days for the Elastic Bunch Graph Matching algorithm on a 1GHz G4 PowerPC processor [36].

Before running these algorithms, the images used need to be preprocessed and require the exact coordinates of the eyes. Normally, this procedure is accomplished manually and the related coordinate data included with each image. The overhead requirement for this step is rarely included in the discussion of performance. With this information, the grayscale images are ready for several preprocessing steps.

The first of these steps is converting the 256 possible integer gray levels into floating point equivalents [36]. The next step aligns all images based on the manually selected eye coordinates in step called geometric normalization [36]. Although not discussed in the documentation, this process allows images to be centered, resized and rotated based on manually extracted data minimizing many of the challenges encountered with scaling and aspect or pose variations. With all images aligned and sized an elliptical mask is used to crop the images to reveal only the face surface from forehead to chin and from cheek to cheek [36]. An example is illustrated in Figure 66.

Figure 66: Preprocessed Image Used in the CSU Face Evaluation System [36]

Taking the aligned and cropped image, the next step, called histogram equalization, more evenly distributes the contrast levels of the remaining face image. Finally pixel normalization is conducted that scales the mean of pixel values to zero and a standard deviation of one. Although not a normal part of algorithm discussions, these preprocessing steps play a very important role in standardizing images and have a significant effect on the overall performance of the face recognition algorithms.

The CSU Face Recognition Evaluation System was strongly considered as a developmental and testing environment due to its popularity and assortment of capabilities and algorithms. However, the related run times would prevent the real-time display of face recognition applications and the existing paradigm of tackling face recognition would be a limitation as more complicated hyperspectral images would be processed, matched and fused.

Instead, it was decided to explore the strengths of the MATLAB® environment and its high-level language. The proven computational ability, numerous toolboxes and extensive library of existing code offered a hope of processing and exploiting the hidden

ability contained within HSI cubes. Among the many attributes were the attractive options of using the image processing and data analysis toolboxes and the accompanied visualization capability. This tool along with the text, *Digital Image Processing Using MATLAB®*, by Gonzalez, Woods, and Eddins, provided an early foundation of knowledge for this research effort [112].

The hardware requirements for this research were met with a laptop and common stand alone desk top. A majority of the initial testing utilized a Dell XPS lap top with an Intel Dual-core T7400 processor (2.16 GHz) with 3 GB RAM and 100GB hard drive. At the end of the research when several algorithms were running simultaneously with the results being combined under different fusion schemes a desk top with Quad–core dual Intel Xeon X5482 processor (3.20 GHz) with 16 GB RAM and a 1 TB hard drive was used for final processing.

With the hardware and computational environment in place, it was time to explore the CMU database. Unfortunately, after inspecting the CMU database it became apparent that some of the original data was missing, probably due to the change over in researchers since the original collection effort several years earlier. A further assessment of the available images revealed that only a total of 48 from the original 54 subjects were available with only a subset of 36 subjects having the necessary minimum of two images, one for the stored gallery and the other as test probe. The subsequent breakdown for the subject numbers with greater than two images available is listed in Table 4.

Carnegie Mellon University Database

- 48 Total Subjects
- 450nm-1090nm (65 bands /10nm increments)
- 640x480 pixels
- 12-15MB per image cube
- 4 light settings per setting (left, center, right, all 3)

16 Subjects have **5 sessions** of data
- Subject #s: 1,2,7,14,18,19,20,21,22,23,24,25,30,31,34,36

6 Subjects have **4 sessions** of data
- Subject #s: 4,5,12,15,26,41

6 Subjects have **3 sessions** of data
- Subject #s: 8,9,10,11,13,28

8 subjects have **2 sessions** of data
- Subject #s: 6,16,17,29,33,37,38,40

12 Subjects have **1 session** of data
- Subject #s: 3,27,32,35,39,42-48

Table 4: Contents of CMU Database

Although each session contains four light settings, all but the setting using all three halogen lamps prove very difficult to work with given the sensitivity of the spectro-polarimetric camera. Denes noted the relative lack of sensitivity by stating that only about 5-10% of the light was useful and the next generation camera would be capable of providing better signal to noise ratio throughout the spectral range of the camera [98]. Denes also found that the darkened and noisy images were "not sufficient to provide adequate discrimination using current face asymmetry algorithms" [98]. A sample of these images under the best lighting conditions is shown in Figure 67.

Figure 67: Grayscale Images before Preprocessing

The initial training set that would form our gallery used the 36 subjects that were available for two sessions. The second session for these 36 subjects were used to form the test set or probes that would be used to assess the accuracy of matching algorithms. A grayscale representation of these image cubes for the gallery are renumbered in order and depicted in Figure 68.

Figure 68: Training Subjects

Adopting the same methodology from the CSU evaluation system, the eye coordinates of the subjects were manually selected and used to accomplish any geometric normalization, followed by cropping the faces with elliptical masks and then performing histogram equalization. Examples of the preprocessed grayscale images are shown in Figure 69, with the resulting training set shown in Figure 70.

Figure 69: Grayscale Images after Preprocessing

Figure 70: Training Images after Preprocessing

Spatial Recognition

After building the preprocessing functions, the image data was ready for initial testing. The first area explored was the spatial domain of the hyperspectral data in the form of grayscale face images. The primary method for image matching was the eigenface method devised by Turk and Pentland [21]. This holistic approach was developed as an attempt to replicate the human recognition process and also as an alternative to many feature based methods that utilized specific attributes but unfortunately discarded much of the surrounding image and contextual information.

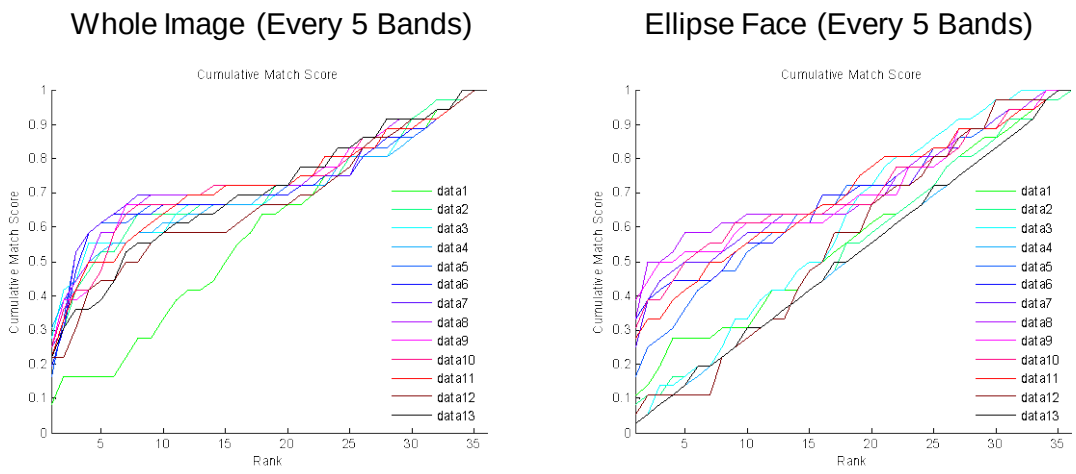

Figure 71: Initial Eigenface Testing Before Preprocessing

For this step, the proven capability of the eigenface algorithm was used with several variations in the preprocessing of the images. A sample of the recognition performance is illustrated in Figure 71 as images from the various frequency bands were tested with the eigenface algorithm. Both cumulative match score (CMS) plots display the range of performance using grayscale images based on only a single wavelength of

the data. The CMS plot on the left does indicate additional performance capability compared to the plot on the right where an ellipse was used to manually crop each of the face images. This additional performance, using the entire image, may be a reflection of the added value of external features mentioned earlier by both the cognitive researchers and the portrait artists.

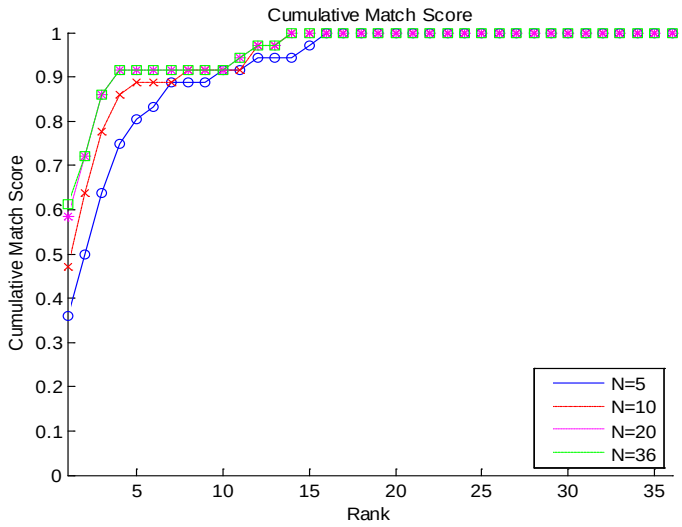

Figure 72: Eigenface Testing - Retaining Varying Number of Eigenfaces

The next element of the eigenface testing was determining the value of utilizing additional eigenfaces. In Figure 72, the previously discussed value of retaining additional eigenfaces in the algorithm is reinforced, as is the point of diminishing returns. Utilizing only 20 eigenfaces from the gallery achieves the same recognition performance as using all 36 eigenfaces. Additional reductions come at the expense of decreased performance.

Similarly, the value of using a combination of wavelengths for spatial information was explored in a simplistic manner. Figure 73 illustrates the potential of combining images from various wavelengths by simply taking the average of all wavelengths and using the mean image for comparison purposes. The small increase in performance is

only an indication as the earlier matching efforts by wavelength showed that some wavelengths were more useful than others were.

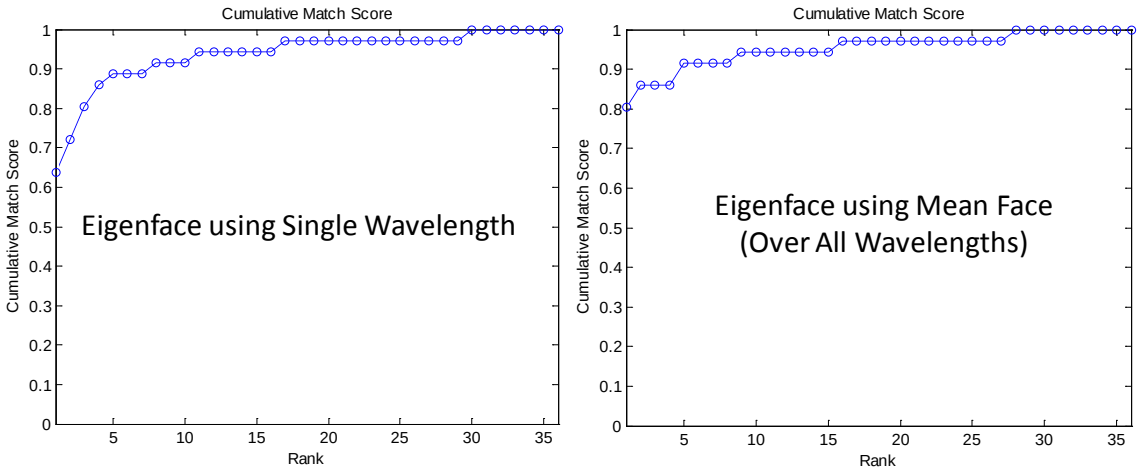

Figure 73: Eigenface Testing - Single Wavelength vs. Average Image

Many variations were tested, including full images, cropped images, resized images using a standard pixel distance between eye coordinates, horizontal leveling of images, varying number of retained eigenvalues, mean images obtained by averaging throughout the frequencies and images with and without histogram equalization. Changes in the overall matching performance were observed but the previous discussed limitations of the image data and this recognition approach became evident. The maximum performance was achieved using a combination of manual preprocessing steps including the simple alignment of eye coordinates, cropping with elliptical facemask and histogram equalization. The results are shown in Figure 74.

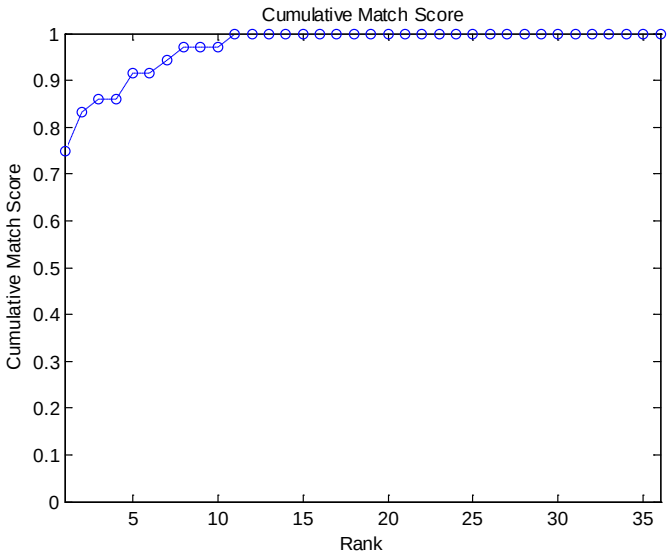

Figure 74: CMS Using Eigenface on Manually Preprocessed Faces

Spectral Recogntion

After the spatial exploration of the hyperspectral data cubes, the next step was to evaluate the matching performance of the data with respect to the spectral dimension. Using manual selection, in a similar manner to Pan's research [96], a sample of 17x17 pixels were selected for each subject from their forehead, chin, and both cheeks. A smaller 9x9-pixel sample was extracted from each subject's lips due to size constraints. An illustration of a subject's sample points are shown in Figure 75.

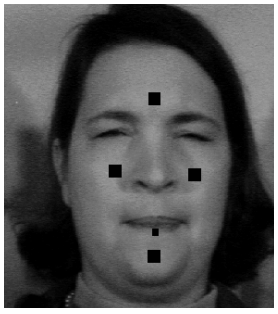

Figure 75: Manual Selection of Points for Spectral Matching

Following methods used by Robila [93], spectral matching capability was evaluated using several variations. The first and most straightforward was by simply using a comparison of the average spectral angle. The variability of the spectral signatures, especially at the sensor wavelength limits, did have an effect on the overall performance. With that in mind, several of the wavelengths at the end of the frequency span were iteratively removed until maximum recognition performance was achieved. The performance limit was still disappointing but never the less expected given the experimental nature of the sensor and resulting data. The results are expressed in a cumulative match score (CMS) plot for each tissue type and are displayed in Figure 76.

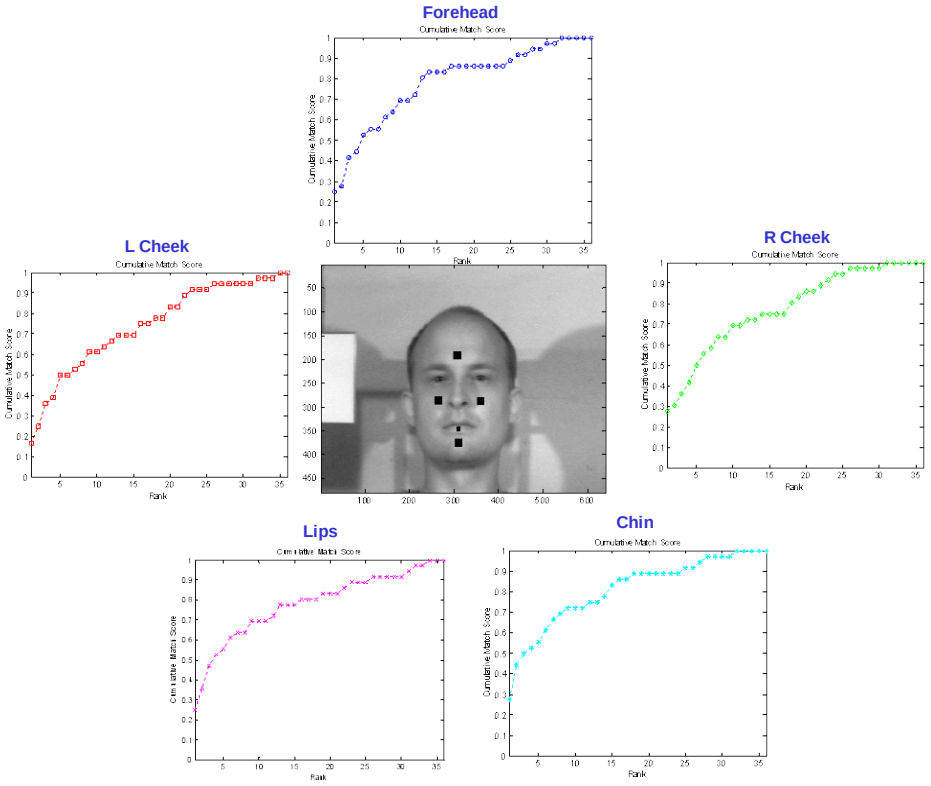

Figure 76: Spectral Angle Matching from Selected Tissue types

In an attempt to incorporate fusion strategies similar to those used by Pan [92], several of the better-known methods were implemented to improve the overall recognition capability. Expecting some correlation among the matching capability, a small improvement was realized. Figure 77 depicts the CMS plots for a combination of tissue samples.

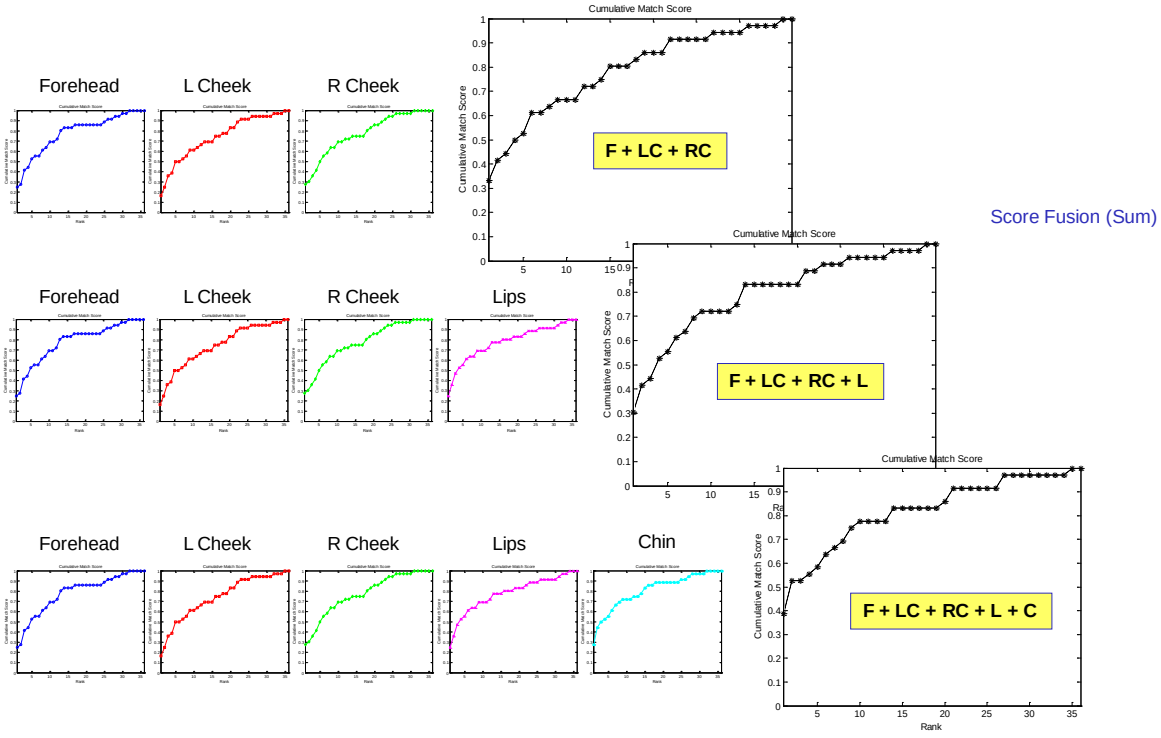

Figure 77: Sum Fusion of Spectral Signatures

In an attempt to improve upon these results, the approach from Elbakary [97] was replicated with the hope of incorporating the noteworthy performance he obtained with the same database. Elbakary [97] used the K-means clustering algorithm to segment skin and then matched subjects based on the Mahalanobis distance measurements. Through iterative testing, it appeared he obtained the best results by designating 4 classes or clusters. An illustration of the results from running his K-means algorithms is shown in

Figure 78. From these designated classes, the selection of the skin class became a little vague.

In Elbakary's words [97], "The cluster among the face clusters that contains more pixels is selected to represent the reference face since more pixels in the cluster implies more information in that cluster." From this confusing description, it appears that the predominate segment is chosen as the skin class for matching. After inspecting the K-means cluster results, identified by the black pixels and shown in Figure 78, some questions remain. If the lower left image contained the largest number of pixels, which is debatable, there still seems to be a significant number of pixels identified in this cluster from the surrounding background. This raises a suspicion that the accuracy of any type of spectral matching would be affected consequently.

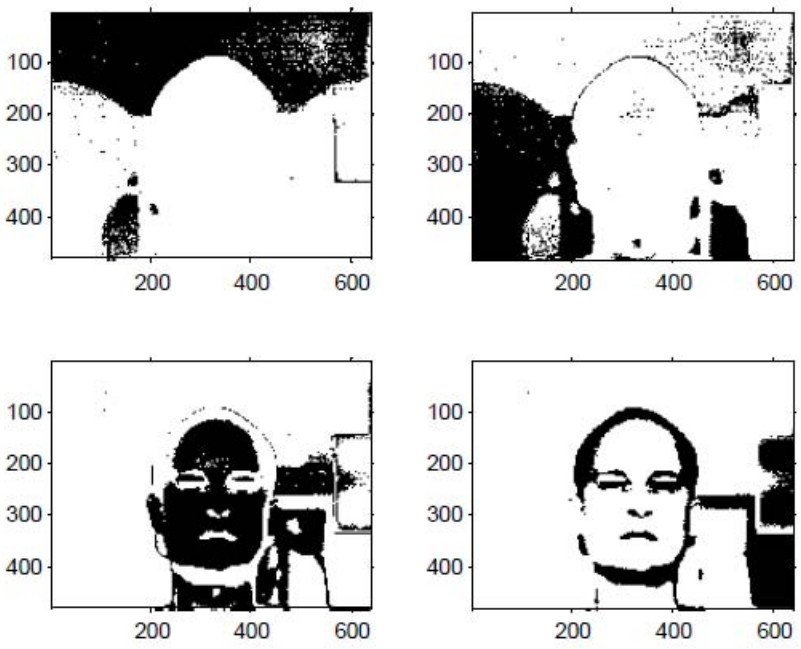

Figure 78: Clusters Shown Separately Using K-means Algorithm [97]

The entire methodology and segmentation of the hyperspectral images were replicated with the single matlab *K-means* function. All image results depicted in Elbakary's work agreed with the segmentation results achieved with the matlab function. The results obtained in Figure 78 are identical to those in Figure 79, as all segments are color coded and combined in a single image, albiet the image in Elbakary's work is flipped from left to right. In this single representation, it is evident that the segmentation of the skin segment is not as accurate as it should be in theory. Given the marked differences in the spectral signatures between live skin and the inert background of the test studio, these results are most probably segmenting on the highest illuminated surfaces of the image. Recall that three 600 watt halogen bulbs were used, and as Denes [98] states, the lamp intensities were "at near the upper end of commercial studio lighting." As such the illumination saturation is probably causing the negative impact on image sementation.

Figure 79: K-means Clustering Results

To gain additional insight and explore the utility of this approach further, additional K-means runs were attempted steadlily incrementing the number of segments. These attempts better segmented the skin from the image background than the previous attempts using 4 segments, but there remained the challenge of devising an automatic selection of the segments associated with the face. Additional excursions using samples from the original database under less than ideal lighting were also evaluated. Again, regardless of the lighting and numbers of segments used an accurate method of identifiying only skin segments or all of the skin segment was not achieved. The results of these tests are shown in Figure 80.

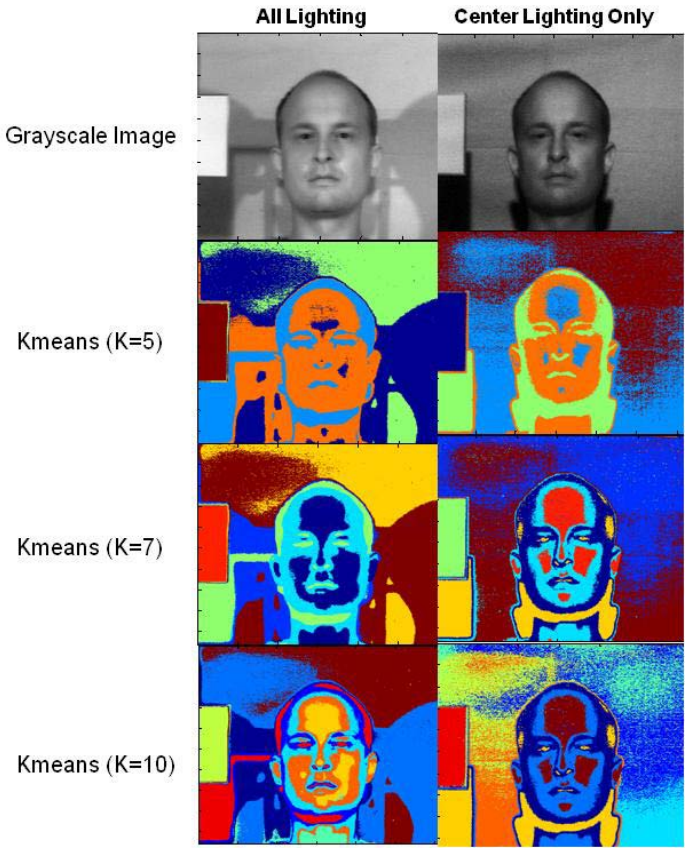

Figure 80: Additional K-Means Testing and Different Illumination

The images were also preprocessed with elliptical masks, exposing only skin surfaces, and were again tested to evaluate the utility and accuracy of K-means algorithms. Unfortunately, the potential contribution to more detailed segmentation would be limited but still a valuable and contributing element in a fused hierarchy.

Unable to refine the segmentation aspect of this approach, the focus turned towards replicating the performance of the method. Elbakary cited impresive results despite the apparent segmentation accuracy and the challenges of the sensor data collected. These results are shown in Figure 81. The statement that these results were comparable to other algorithms in the literature is true as Pan's [58] results are displayed in the figure that follows (Figure 82). Pan's results were achieved with a similar Mahalanobis distance comparison but was matched against the manual selection of skin pixels from a higher quality database [58].

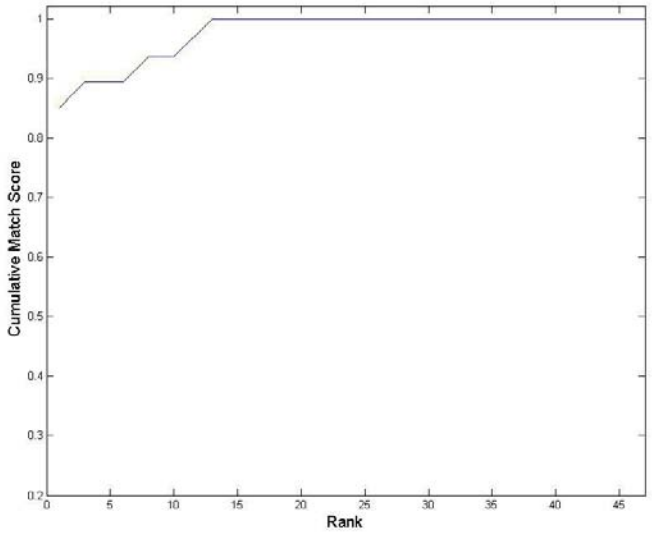

Figure 81: Elbakary's Published Results [97]

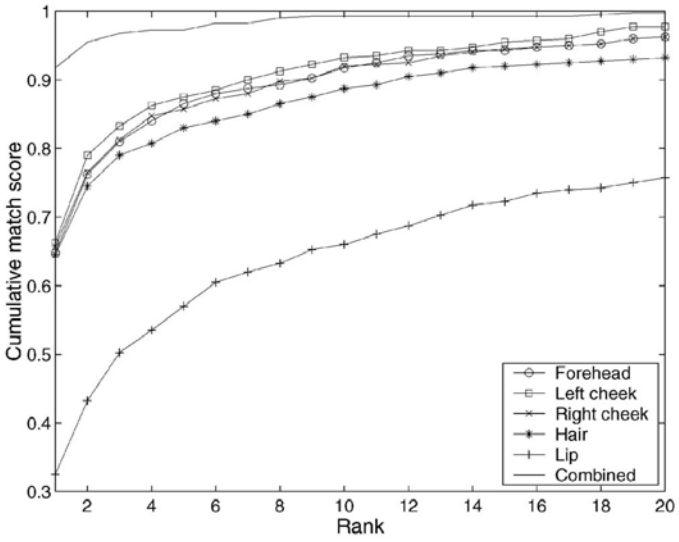

Figure 82: Pan's Published Results Using Spectral Signatures [58]

Surprisingly, Elbakary matched Pan's level of performance utilizing data that was obtained with an inferior sensor and observable inaccuracies in the segmentation results. Motivated to match these results, additional adjustments were made to increase the effectiveness of the spectral matching approach. Multiple experiments were used to refine or filter the skin samples obtained from the K-means segmentation. Using filtering thresholds based on standard deviation, outlier detection methods, and simply removing the noisiest frequency bands at the end of the sensor performance limits were used to increase the matching accuracy. Also, supplementary comparison methods besides mahalnobis distance were are utilized to include spectral angle and gradient matching. Although some additional performance gains were achieved, the overall levels hoped for were never replicated.

On the otherhand, the by product of these efforts was a better understanding of the performance limits and computational cost of utilizing these methods. Several of the resulting CMS plots are illustrated in Figure 83. The first plot (top left) displays the intial

performance obtained using the Mahalanobis distance comparions Elbakary employed followed by the addition of outlier filtering (top right) using the blocked adaptive computationally efficient outlier nominator (BACON) algorithm from Billor [113]. The bottom two plots reflect an effort to conserve computational effort by using the straightforward spectral angle matching employed by Robila [93]. Even with the removal of select wavelengths near the bounds of the camera's capability, the overall improvement failed to match previously documented recognition capability.

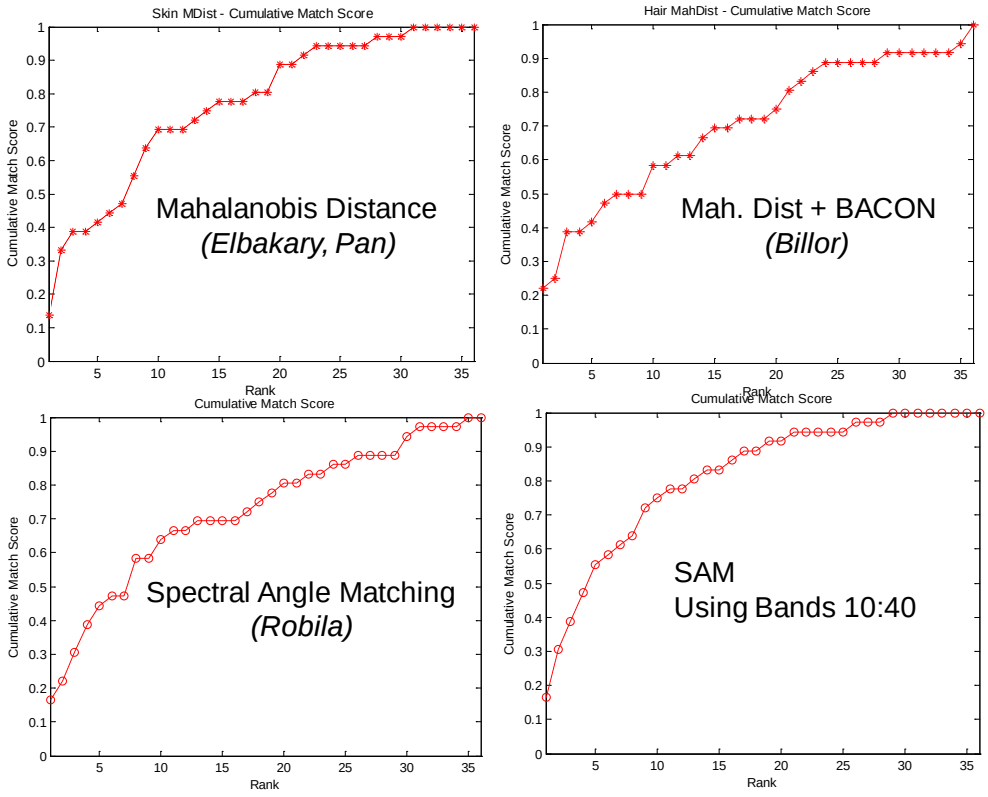

Figure 83: K-means and Spectral Matching Testing

These efforts helped evaluate various methods that could be implemented in both locating skin segments and the subsequent matching strategies. In an ideal world, we

would like the best performing classifier possible, but when combining classifiers in an ensemble the diversity of these approaches can prove to be a benficial ally. Kuncheva [86] states that "even weakening the individual classifiers for the sake of better diversity appears to be an excellent ensemble building strategy." This is possible only if the classifiers make different types of errors that are complimentary. This diversity can measured in many ways as Kuncheva [86] points out with relationship measures of simple correlation, Q-statistic, interrater aagreement, pairwise (disagreement and double fault measures) and non-pairwise measures (entropy and Kohav-Wolpert Variance). With these evaluation techniques the design of any ensemble can be carefully adjusted for diversity and ultimately maximum performance.

Although the results were not replicable, the approach still proved to be a useful method and reinforced an ability to identify individuals simply by the spectral signature of their skin, reinforcing the findings of Robila and Pan. This spectral testing was also beneficial in indentifying some weaknesses of the published approaches and offered some clues for improving and integrating this capability with other methods.

Interest Point Recognition

The next area of exploration was testing Lowe's SIFT method which exploits invariant features for object identification [60]. SIFT extracts these features or key interest points using a Difference of Gaussians function. The local minimum and maximum of this function are used to create a feature vectors that describes the orientation and gradient based on neighboring pixels. These features are shown to be

invariant to image scaling, translation and rotation. These interest points and matches are shown for other face images under normal conditions, rotation, and scaling in Figure 84. For our initial test subject, SIFT proved just as capable of identifying faces as it did for object recognition, even when an artificial rotation was incorporated into the face image.

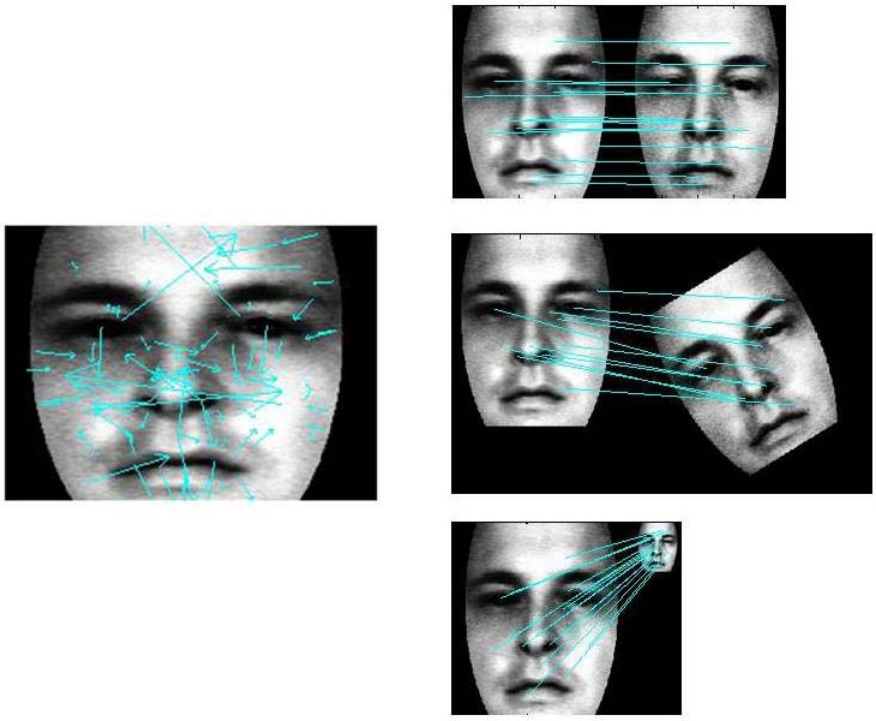

Figure 84: Application of SIFT to Determine Interest Points and Matches

Despite the promise of this application, the implementation of Lowe's executable code proved to be the most impressive performing algorithm but also the most computationally intensive. Several excursions using the SIFT algorithm were made with incremental distance thresholds for the variable distRatio, which is the ratio of Euclidean distances between the nearest and the second nearest neighbor. The results, shown in Figure 85, illustrate the effectiveness of the algorithm.

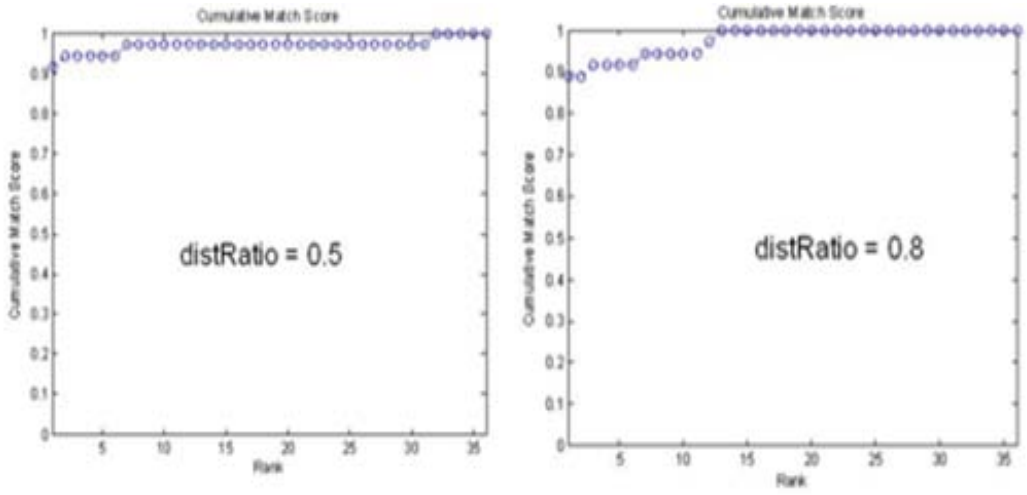

Figure 85: SIFT Experimental Results with Varying Distance Thresholds

In general, there was minimal impact of this variable as the same subgroup of four subjects proved to be difficult to match. These subjects and their images from the training and test data sets are shown in Figure 86. By inspection, it appears that the squinting and hair coverage, for the first subject, and addition of glasses, for the last three subjects, causes trouble for the SIFT algorithm which is focused on matching the internal features of the face. Recalling the results from cognitive research, it is the combination of both internal and external features that provides that provided the best performance when it came to human recognition capability. Extending this lesson to our growing suite of recognition algorithms, perhaps the combination of the holistic eigenface algorithm and the complimentary interest point algorithm would mimic this approach and capability.

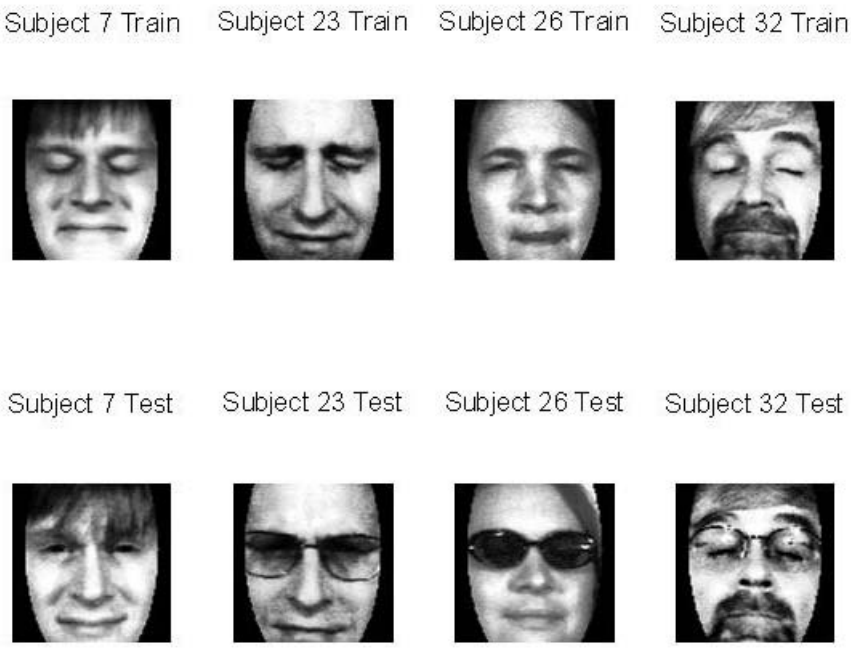

Figure 86: SIFT Recognition Errors

As mentioned earlier, the SIFT algorithm came with a computational drawback. Although much of the copyrighted code (available free of charge to academic research efforts) comes in excutable format, the processing time to run grayscale images for a data set of only 36 subjects could exceed an hour on our computer laptop. To effectively employ this capability, a reduction in processing time would become nessary. Insight on how to accomplish this was provided by looking at earlier results compared to those after data reduction efforts that would occur as a result of the preprocessing images. Using the initial images it was not uncommon to experience run times in excess of an hour.

However, by discarding a significant amount of the irrelevant data and reducing the size if the input images a considerable amount of processing time would be saved and in this case performance would increase. As an example, the initial 480x640 pixel image

required 68 minutes to process the entire database. When a simple square template was placed on these images, reducing the effective size to 251x251 pixels, the processing time was cut in third to 21 minutes. The fortunate byproduct of this reduction in image size was not only processing time but also an increase in recognition performance.

Published Algorithm Comparison

After testing some algorithms on the spatial and spectral domain of the HSI, the next logical step was to integrate the two domains into one approach. This is not entirely a novel approach, as Pan implemented a clever approach called a spectral face in an earlier research effort [75]. In this effort a face representation was created by assembling an image by incrementally selecting each subsequent pixel from the following band and repeating throughout the image. Replicating this effort to capture the spectral and spatial information into one image, the CMU database was transformed into a similar format and is shown in Figure 87, next to a comparable image from the CAL database used in Pan's research.

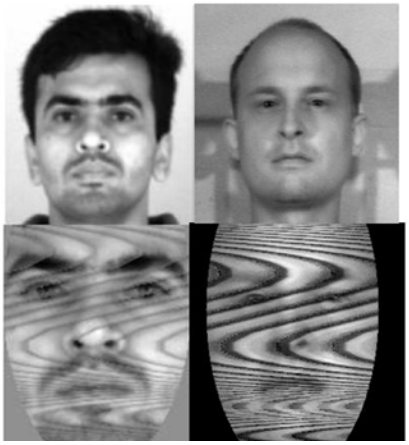

Figure 87: Spectral Face Implementation for CAL and CMU Databases

Although the method and images are similar, the results were unfortunately not. The continued trend of below average performance for all tested algorithms simply confirmed the troublesome nature of the CMU data. This is only one perspective however. An operational employment of a recogntion system will seldom experience the pristine conditions offered by the AT&T [119] and CAL [96] databases as variances in pose, lighting and sensors frequently provide challenging real world data. The results of replicating some of the leading algorithms has fallen short of published results but the value of these approaches is not lost. A comparison of the overall performance is depicted in Figure 88.

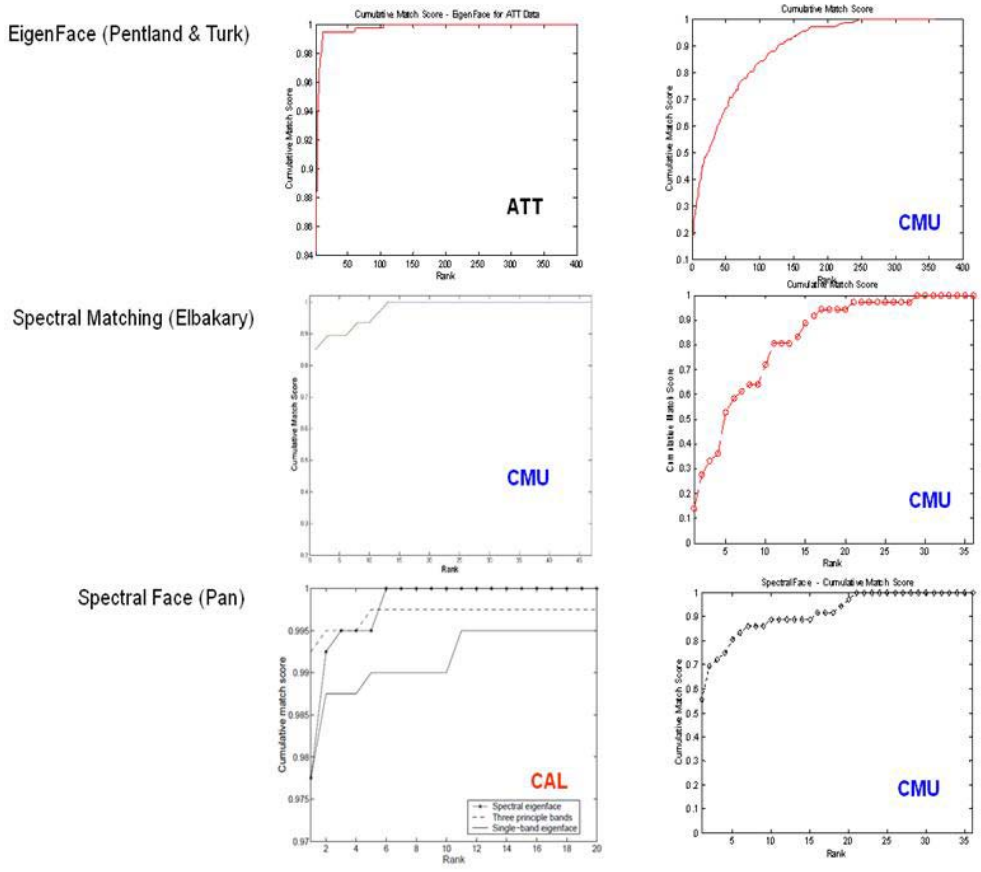

Figure 88: Testing Results of Select Algorithms on CMU Data

Findings from this initial round of testing suggested that a fusion framework would need to combine complimentary aspects of these algorithms to provide the required performance capability regardless of data quality or environmental setting. Taking into account the processing time of some of the algorithms, a method to incorporate data reduction wherever possible should be devised to reduce the overall computational time. The next section will describe both the efforts to improve the individual algorithms and the attempts to integrate them into a hierarchy for a robust face recognition system.

Algorithm Enhancements

In early testing the generally accepted preprocessing practices were used, but as the computational capability increased, it seemed unacceptable to accept these artificial methods that would restrain any operational implementation of a system. The first step to adjust this process was incorporating an automatic segmentation of the skin and hair. As shown earlier, with the testing of Elbakary's K-means clustering, the identification of skin surface was possible. The use of this approach did not appear to be easily adjusted for an automated approach because of the problems with accuracy and computational speed.

Nunez's [55], [68] research however provided a method with NDSI to identify skin surfaces using only two wavelengths from hyperspectral images. The technique and reduction offered an attractive option to a more involved clustering method. Unfortunately, NDSI looked for two key wavelengths 1080nm and 1580nm in order to

calculate the index. The CMU data only spanned the spectral range from 450nm to 1090nm. So only one of the key wavelengths was contained in the data and the one wavelength included was located at the performance boundary of the Spectro-Polarimetric camera. Faced with this difficulty, an alternative to this proven method was explored to demonstrate the utility of this approach.

With the advice from Nunez, a less effective, but suitable, alternative was devised that used a combination of indexes designed to highlight the unique characteristics of the spectral signature of human skin and eliminate common confusers. Keep in mind that this work around was meant to mimic the proven method devised by Nunez's research and would be applicable with a more current hyperspectral database.

$$\text{NDSI} \quad \gamma_i = \frac{\hat{\rho}_i(1080nm) - \hat{\rho}_i(1580nm)}{\hat{\rho}_i(1080nm) + \hat{\rho}_i(1580nm)}$$

$$\gamma_{Vi} = \frac{\sum_{j=1050nm}^{1080nm} \hat{\rho}_i(j)}{\sum_{k=450nm}^{480nm} \hat{\rho}_i(k)} \qquad \gamma_{Wi} = \frac{\sum_{j=890nm}^{910nm} \hat{\rho}_i(j)}{\sum_{k=970nm}^{990nm} \hat{\rho}_i(k)}$$

$$\gamma_{Xi} = \frac{\sum_{j=1060nm}^{1080nm} \hat{\rho}_i(j)}{\sum_{k=970nm}^{990nm} \hat{\rho}_i(k)} \qquad \gamma_{Yi} = \frac{\sum_{600nm}^{620nm} \hat{\rho}_i(j)}{\sum_{540nm}^{560nm} \hat{\rho}_i(k)}$$

Equation 30: NDSI Substitute Approach

By examining the available wavelengths in the data as well as the quality of the information, an alternative approach was designed to sum relevant wavelengths and create indexes similar to NDSI that exploited the spectral characteristics of skin. Seen

below in Equation 30 is the NDSI calculation and the alternative indices. Below the NDSI equation are four indices used to highlight the increase in reflectivity in the NIR wavelengths versus blue wavelengths (γ_{Vi}), highlighting the characteristic water absorption dip at 980nm $(\gamma_{Wi},\gamma_{Xi})$ and a final check to remove potential plant material that can act as a confuser (γ_{Yi}).

By combining these indices that indicate the possibility of skin, when the value is greater than one, the skin segment can be identified rather efficiently compared to K-means. The most effective implementation of this approach relied on (γ_{Xi}) and (γ_{Yi}) indicators to identify skin potential pixels. All pixels in the hyperspectral image cube that fell near the calculated average of the potential skin pixels were deemed a skin surface. Additional image processing methods were employed to filter noisy pixels. Among these included the matlab function, *graythresh* that employs the Otsu method that chooses the threshold to minimize the intraclass variance of the black and white pixels and the matlab function, *bwareaopen* to remove small objects (pixels).

A similar approach was used for the identification of hair segments in the image. This time using a NDVI calculation (Equation 31) and then fine-tuning the selected segment using Mahalanobis distance comparison for only the red (650nm), green (510nm), blue (475nm) and NIR (1000nm) wavelengths the hair segments, including facial hair, was obtained.

$$\text{NDVI} \quad \gamma_i = \frac{\hat{\rho}_i(1000\text{nm}) - \hat{\rho}_i(650\text{nm})}{\hat{\rho}_i(1000\text{nm}) + \hat{\rho}_i(650\text{nm})}$$

Equation 31: NDVI Calculation

These procedures are not as straightforward as initially intended but with the appropriate wavelengths, the approach can be simplified. The results of these two segmentation steps for skin and hair are illustrated in Figure 89. The implication of this capability would soon be exploited in subsequent steps.

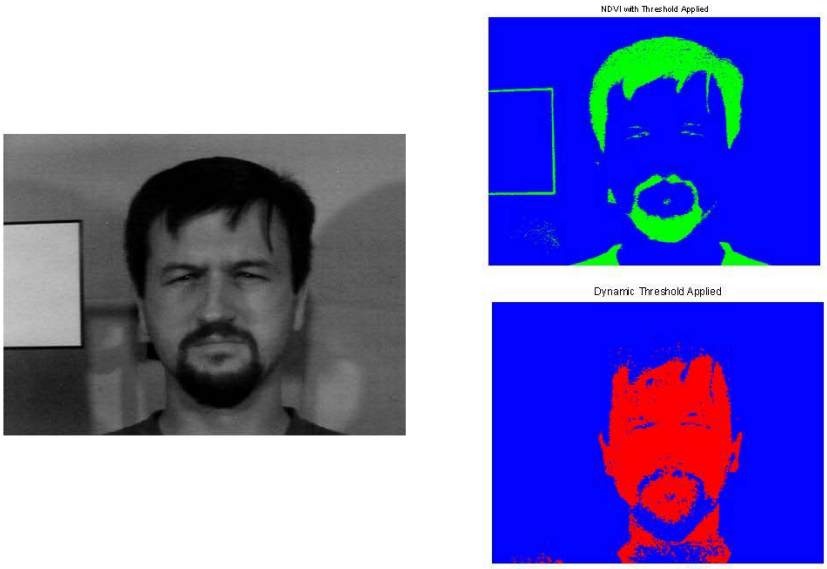

Figure 89: Contextual Layers of Skin and Hair

With the unique ability to segment the skin and hair segments of the image, it was uncomplicated to include a centroid calculation to accomplish the task of automatically centering images for identification. The additional steps that are routinely accomplished during preprocessing could now be enabled automatically. These adjustments include the centering of all face images, leveling in the case of unintended rotation of the face, resizing the image for a consistent scale across individuals or the population. Once this is accomplished, the removal of background clutter is accomplished by the application of an elliptical mask. Unfortunately, when this is accomplished some important information is removed from the image including the relative shape of the head and a good portion of

the hair on top of the head. This same approach was initially attempted but as our processing capability matured we found this step crude in its application.

Additionally, the requirement to provide two eye coordinates for size adjustment proved unnecessary as other measurements could be extracted from portioned skin and hair segments. If desired, standard image sizes could be automatically obtained by scaling measurements from the segmented face such as the major axis, minor axis and an eccentricity. Not only were these measurements helpful in automating a previously laborious manual process, but now provided additional descriptive feature that could be incorporated into the recognition system. The centroid location is calculated and depicted in Figure 90.

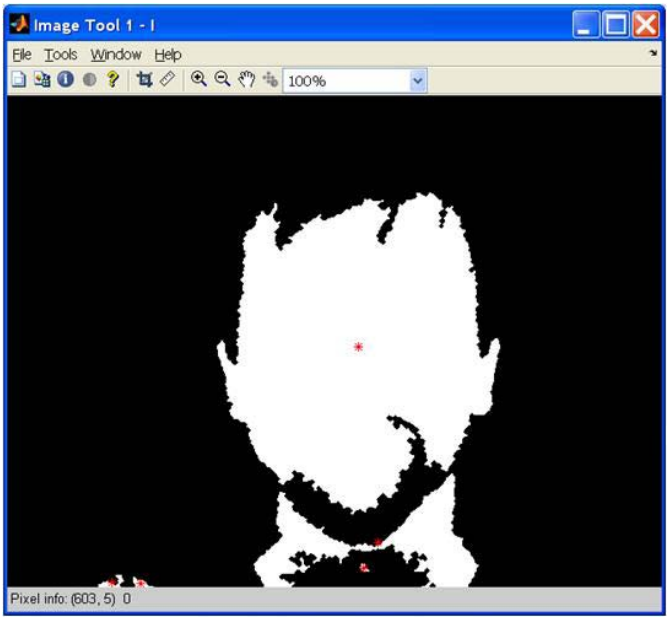

Figure 90: Using Skin Detection to Determine Centroid

With the detailed segmentation, the face and hair segments could be easily extracted from any image while keeping valuable external cues such as shape and hairstyle that would be used later. Finally, these cropped images are next passed through a histogram equalization adjustment to utilize the full range of grayscale intensities and to bring out some of the less noticeable features.

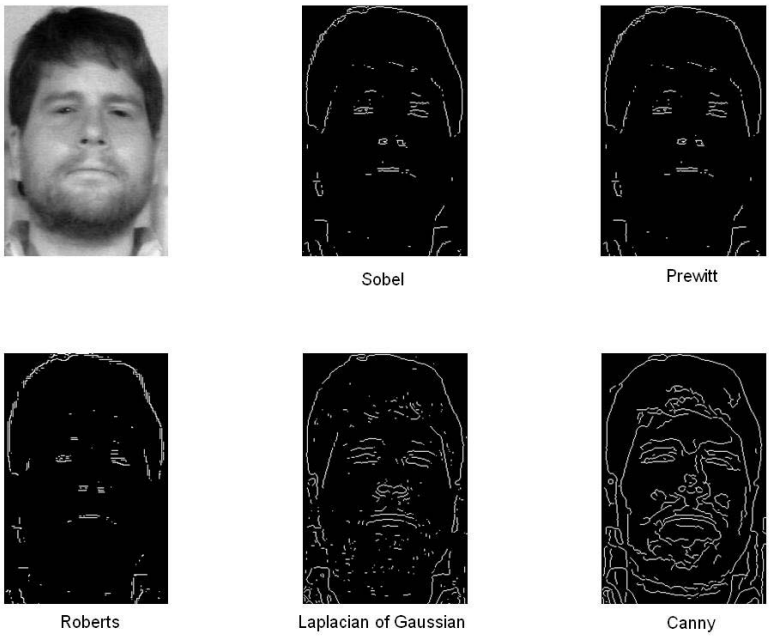

Figure 91: Edge Detection Efforts

With the ability to segment and collect features related to the external aspects of the face the next endeavor was to capture more detailed information contained in the high frequency data of the face. Additional efforts to accurately collect this data included several proven edge detection methods included in the matlab image processing toolbox such as the Sobel, Prewitt, Roberts, Laplacian of Gaussian and Canny methods. The most effective of these methods was the Canny function that locates edges by finding the local

maxima of gradients in the image. This particular method worked best because it employs two thresholds, one for strong edges and the other for weak edges. Weak edges are only included if they are in proximity to strong edges making the method very useful in noisy data. The results of these methods are shown in Figure 91. It was hoped that these features would provide a useful layer of contextual information, since both low and high frequency information make up two important components of two-dimensional images. Unfortunately, matching results incorporating this information into a composite image, shown in Figure 92, was discouraging and prompted further exploration using other approaches that would add measurable value to our fusion strategy.

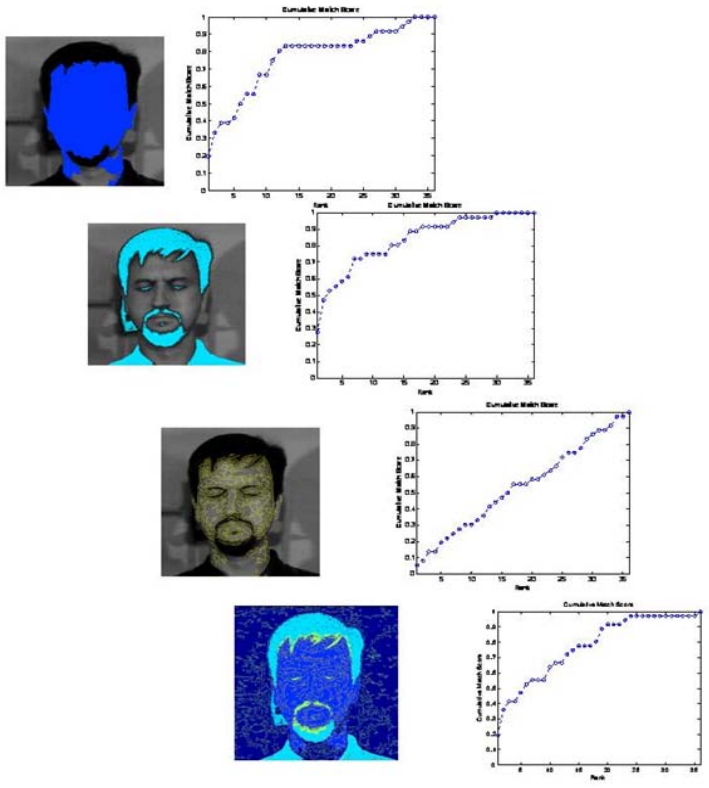

Figure 92: Incorporation of Edge Information for Recognition

In Pan's research, he showed the value of adding distances obtained as each principal component of a hyperspectral image cube was matched with other images. Additional information was observable in each of these components and at the same time the span of data was efficiently reduced as the many wavelengths could be reduced to several components that contained a majority of the information contained in the data. Investigation into this method for our data was attractive for the same reasons. Unfortunately, once the results were obtained, the reoccurring reminder of quality issues for the CMU data was evident.

Figure 93: First Five Principal Components: CMU Data (Top) and CAL Data (Bottom)

Only a glimpse of the principal components of the CMU and CAL data is required to see the difference between the data sets. Recall that the CAL data spans only a spectrum subset (700nm–1000nm) of the CMU data (450nm–1100nm) with about half the number of bands (31 versus 65). Despite this apparent advantage, the evaluation of the principal components displays an inferior set of information. When viewing the CMU components only a few images appear to have visible detail where the CAL data

displays unique and complimentary detail (Figure 93). It is no surprise, that when Pan [75] utilizes a Mahalanobis Cosine distance comparison to match sequential principal components, the additive scores provide improved accuracy that can be incorporated into his recognition testing.

For spectral signature matching, the addition of outlier detection was investigated using the blocked BACON algorithm [113]. Smetek et al. [121] conducted tests on multiple multivariate outlier detection methods on hyperspectral images and found that the BACON algorithm was advantageous because of its computation speed as well as its low number of false alarms. Initially, it was hoped that this ability to detect outlying spectral signatures would enable us to identify unique skin features such as blemishes, moles, or freckles that are caused by anomalies in the epidermis. Unfortunately, despite the theoretical promise of this approach, experimental testing of this capability was unsuccessful, perhaps in part because of the data resolution and quality issues discussed previously. Using this insight, we chose the BACON algorithm amongst several available to look for improvement in refining skin and hair samples and ultimately performance. A quick overview of Billor's outlier detection method follows.

Most statistical methods utilized on multivariate data assume homogeneous data and rely on robust investigative methods to relax this assumption. This assumption is complicated by the fact that real data is routinely comprised of imperfect, partial, and missing information. Unfortunately, there are consequences of this direct approach, as discrepancies or outliers impact analysis through covariance distortion and breakdown. Surprisingly, these effects can occur with as little as the influence of one outlier and with the methods and data being used in this research, outliers were very likely.

One of the first challenges in developing an outlier detection method is creating a metric that will not be contaminated by the very inhomogeneities that are hidden in the data. Detection algorithms often abandon optimality conditions and work through an iterative process to locate outlier candidates. A result of this iterative process is an unattractive algorithm due to its computational expense. This expense quickly escalates with increasing sample size to the point that sophisticated approaches become infeasible on large databases. Certainly, for this application, this was a real concern.

Even effective and efficient detection methods face pitfalls that must be avoided during the investigative process. The first effect is known as masking and occurs when covariance inflation appears due to the effect of strong outliers that "mask" the presence of outliers [121]. Examples of this are ellipsoids, defined by a distance threshold that grow to include outliers in the dataset. Many detection methods can also be influenced by an effect known as swamping. The swamping effect denotes an increase in an outlier detector's false alarm rate that can occur when the covariance ellipsoid is either shifted or rotated so that "good" observations no longer lie within the threshold ellipsoid or within a designated boundary [121].

To better understand and discuss the detection algorithm it is necessary to review some basic terminology and characteristics. The algorithm is applied to datasets of n observations, p variables, and a number of outliers equal or less than half the number of observations ($k < n/2$). A desired characteristic of this algorithm is the ability to function even when data is highly contaminated, again a high potential given the observations in earlier data exploration. This attribute is referred to as a high breakdown point and is better described as the smallest portion of the dataset that can be modified without

making the estimator unreliable [113]. A second desired characteristic for a detection algorithm is to be affine equivariant. This means that the identification of an outlier is not dependent on the location, scale, or orientation of the data being examined. In mathematical terms, an estimator, T, is affine equivariant if and only if, $T(\boldsymbol{XA} + \boldsymbol{b}) = T(\boldsymbol{X})\boldsymbol{A} + \boldsymbol{b}$, for any vector $\boldsymbol{b}$ and nonsingular matrix $\boldsymbol{A}$ [113].

The most rudimentary detection approach would be to check all possible subsets of the data (size $k = 1,\ldots,n/2$) and determine whether a given subset is outlying relative to the remaining observations in the data. Although thorough, this approach becomes computationally impractical except on small datasets. An alternative method involves forming a clean subset of the data and then testing the remaining points or subsets. Computationally efficiency can be accomplished if groups of data, instead of single points, are allowed to enter the clean subset during subsequent searches of the data. One example of an alternative approach is known as the Minimum Volume Ellipsoid (MVE) method. In this approach an ellipsoid of minimum volume is defined that contains at least a "half-sample" ($h = [(n + p + 1)/2]$) of observations [121]. Another commonly used method is the Minimum Covariance Determinant (MCD). In this method, a subset of h observations is identified whose covariance matrix has the minimum determinant [121]. Despite the effectiveness of these approaches, they come at the expense of computational cost.

The BACON algorithm provides an iterative, yet efficient, approach to nominate potential outliers in a dataset [113]. In testing the BACON algorithm is able to evaluate and nominate outliers in less than five passes through the data. This approach starts by forming a basic subset ($\boldsymbol{X_b}$) of m observations free from outliers. In this case $m = cp$,

where c is a small scalar (typically 4 or 5 based simulation trials), *p* variables. $\boldsymbol{X_b}$ can be determined by one of two methods. Version 1 builds the clean subset with the *m* observations having the smallest Mahalanobis distance. The Mahalanobis distance is calculated as shown below with $\bar{\boldsymbol{x}}$ and $\boldsymbol{S}$ representing the mean and covariance matrix of the n observations present in the data.

$$d_i(\bar{x}, S) = \sqrt{(x_i - \bar{x})^T S^{-1} (x_i - \bar{x})} \qquad i = 1, \ldots, n$$

Equation 32: Mahalanobis Distance [113]

The approach using the Mahalanobis distance is attractive because it is an affine equivariant approach that offers low computational cost. The drawback of this approach is that it possesses a lower breakdown point (20%) than may be desired [113].

The second method of forming an initial clean subset employs a calculation of median distances, where $\boldsymbol{m}$ is a vector containing the coordinate medians ($\|x_i - m\|$). In this case, $\boldsymbol{X_b}$ includes *m* observations with the smallest distances from the medians. The median distance method offers us a more robust starting subset and a resulting higher breakdown point (40%) compared to the Mahalanobis distance method. The weakness of this approach is that it is not affine equivariant. Once this initial basic subset is chosen we can proceed with the BACON algorithm that is summarized next.

BACON Algorithm for Multivariate Data [113]

1. Select an initial basic subset of observations using Version1 or Version2.

2. Compute discrepancies using Mahalanobis distance method using $\boldsymbol{x_b}$ and $\boldsymbol{S_b}$ which are the mean and covariance of the clean initial basic subset. The formula is shown below.

$$d_i(\overline{x_b}, S_b) = \sqrt{(x_i - \overline{x_b})^T S_b^{-1} (x_i - \overline{x_b})} \quad i = 1, ..., n$$

Equation 33: Mahalanobis Distance of Basic Subset [113]

3. Set new basic subset to all points with discrepancy less than $c_{npr}\chi_{p,\alpha/n}$ where $\chi^2_{p,\alpha}$ is the $1-\alpha$ percentile of the chi-square distribution with *p* degrees of freedom, *r* is the size of the current basic subset, and c_{npr} is a correction factor shown below. c_{hr} is a variance inflation factor when *r* is much smaller than *h*.

$$c_{np} = 1 + \frac{p+1}{n-p} + \frac{1}{n-h-p} = 1 + \frac{p+1}{n-p} + \frac{2}{n-1-3p}$$
$$c_{hr} = \max\{0, (h-r)/(h+r)\}$$
$$c_{npr} = c_{np} + c_{hr}$$

Equation 34: Formulation of Correction Factor [113]

4. Stopping Rule – Iterate Steps 2 and 3 until the size of the basic subset no longer changes. The observations excluded by the final basic subset as are offered as outlier candidates.

With the ability to iteratively add groups of clean observations, an efficiency is gained that allows this method to be applied to datasets as large as one hundred thousand

observations. The improved forward selection method saves computational costs by reducing the number of covariance matrices computed and inverted.

The application of the BACON algorithm proved to be more effective when it came to detecting outlying hair samples. Irregularities were observed in the spectral signatures extracted from the hair perhaps due to varying thickness, location, hair treatments, or just the detection method. A visual depiction of these spectral signatures is shown in Figure 94 before and after outlier detection and removal.

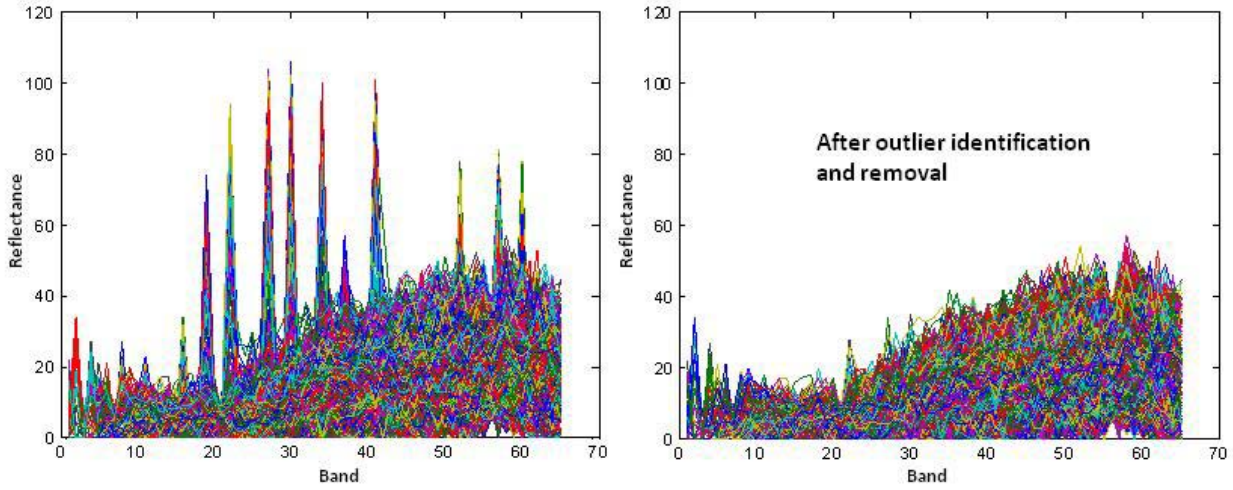

Figure 94: Application of BACON Outlier Detection for Hair Signatures

Despite the apparent success of detecting and removing outlying pixels the overall improvement in matching capability was modest. In Figure 95, the CMS of using Mahalanobis distance is shown before outlier detection and after outlier removal. The shift in performance is positive but does not fulfill the promise hoped for. Given these results, outlier detection methods would only provide a fine-tuning of tissue detection and

matching capability but would not provide the needed improvement necessary for a solitary solution for identification.

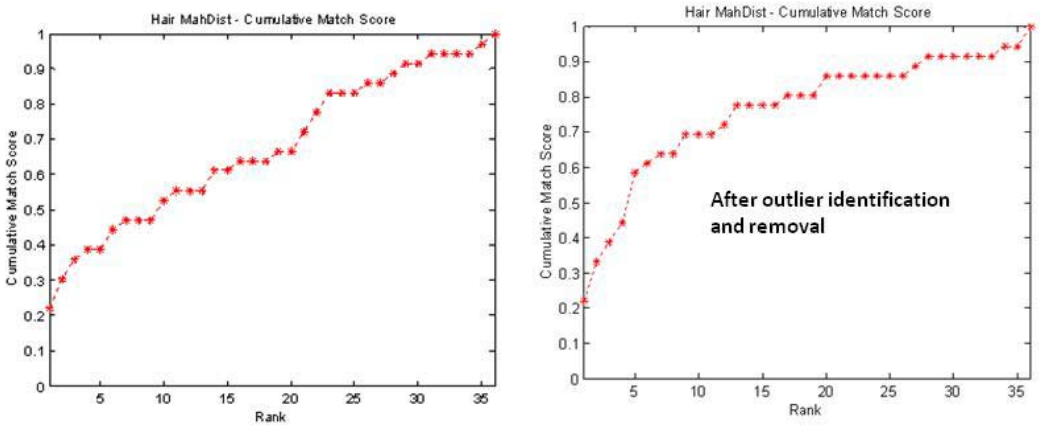

Figure 95: Performance Impact of Applying Outlier Identification

Remembering some of the earlier successful application of NNs for face recogntion such as Pan [26] and Er [27], we hoped to apply the same performance ability to the challenging task of providing insight to the limited success of our spectral signature matching attempts. The earlier testing results indicate that this indicator could serve as a soft biometric cue rather than a tool for identification. This capability could automatically identify skin types, much the same way Jain [38] did with gender, height, weight, ethnicity, skin, hair, and eye color in his application of soft biometric data to a fingerprint matching system.

In Jain's research these various soft biometric indicators were shown to improve recogntion and the performance increase achieved is displayed by his results shown in Figure 96. Although this research used database information as the soft biometric input, Jain hoped that in future efforts he would employ a method to automatically extract these

same soft biometric features [38]. Our short term goal is to see if Jain's goal could be realized through this research.

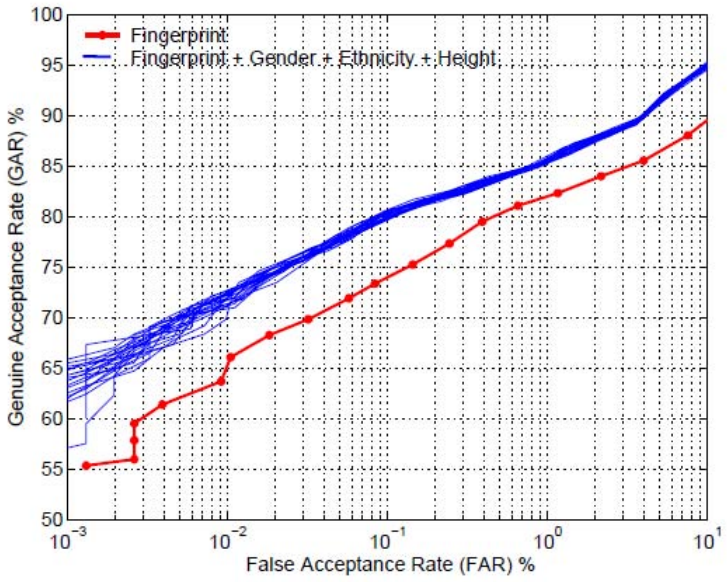

Figure 96: Improvement in Biometric System Using Soft Biometrics [38]

Since the spectral signatures of the skin did not produce the promised performance to accurately identify individuals, perhaps another use that would prove beneficial to the overall recogntion effort. What if spectral signatures could simply be used to extract a characterization or soft biometric typically used in a personal description? The investigation began with the selection a probabilistic NN (PNN) to explore whether pigment types could be consistently characterized. A PNN was selected because of its straightforward application and its ability to generalize with a small data sample.

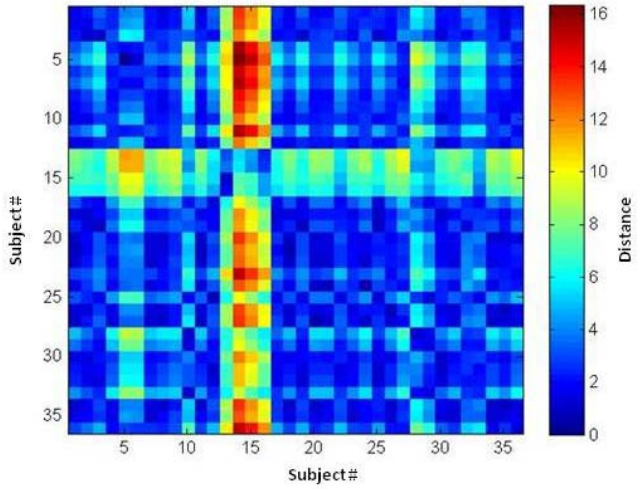

Figure 97: Spectral Angle Distances

A simple inspection of the average spectral angles between subjects indicated that a subset of the group (subjects 13, 14, 15, and 16) had spectral signatures furthest from the other subjects and would be logical candidates for the development of a simple descriptor. The goal of this experiment was to take all odd numbered subjects and the associated target matrix identifying the subjects of interest, 13 and 15, build a NN that could classify this type of skin based on the spectral signature. After building and simulating the PNN in matlab, it was capable of classifying the nearest similar skin signatures, subjects 14 and 16, from the even numbered subjects. This focus group was the most easily designated from the entire gallery as seen earlier by their spectral angle measurements but application to other groups was now feasible. Using the following procedure, the expansion of this application to hair color using the spectral signature is direct but there are also other ways this proven methodology could be applied to extract soft biometrics .

Many of the necessary features obtained with the segmented portions of the hyperspectral face image provide useful descriptive elements of an individual. The segmentation of skin, and the resulting collection of face measurements, presents features that easily lend themselves to a characterization of a human face such as large, skinny, or wide, and with a measure of eccentricity an indication of an odd-looking face. Now, in a similar manner to the way Jain wanted to inject soft biometrics, size, hair, and ethnicity can be extracted automatically and available for incorporation into a recognition system. This automated use of spatial and spectral features can leverage the value and speed of a PNN to classify these various features into descriptive features.

The data exploration effort provided an understanding of the quality of the CMU database. Subsequent development and testing of leading algorithms illustrated the challenges of applying each method independently. Even after further refinement, algorithms focused on both the spatial and spectral dimensions of the data did not provide the necessary performance. In order to achieve higher levels of accuracy and speed a more sophisticated method had to be devised. Common fusion approaches offered some hope but an adaptable and scalable architecture would be necessary if this academic effort were ever to lend itself to an operational implementation for hyperspectral face recognition.

Hierarchical Architecture

From the field of automatic target recognition, Ando provides a useful hierarchy for processing the hyperspectral face images [109]. At the lowest level, processing

includes smoothing and segmenting the image. During mid-level processing, cues such as shading, texture, reflectance, and illumination are extracted. Lastly, high level processing integrates invariant information across different viewpoints for final identification.

Varshney [104] describes the fusion model used in the United States military community, established by the Joint Directors of Laboratories Data Fusion Working Group, shown in Figure 98. This framework allows a standardized understanding and discussion of data fusion issues. Starting with the first two stages, incoming data from sensors (***information sources***) is screened and allocated (***source preprocessing***) based on priority and throughput limitations [104]. The ***three levels of processing*** progress from locating and tracking entities, to assessing relationships between objects, and finally making inferences about the current situation [104].

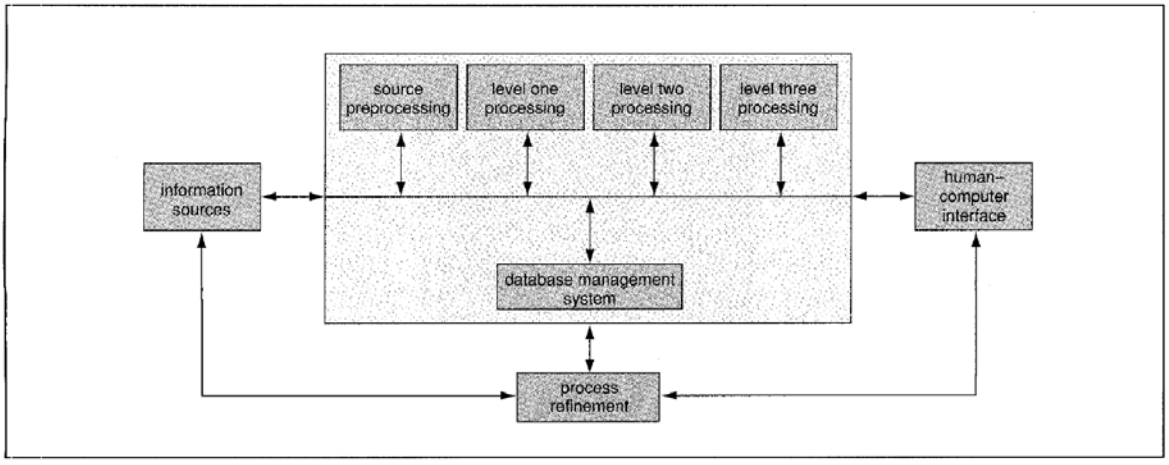

Figure 98: Joint Directors of Laboratories Data Fusion Model [104]

Process refinement is a meta-process that monitors performance and makes recommendations for improvements [104]. ***Database management*** is concerned with

managing the retrieval of relevant data and the associated requests for related information [104]. Finally, the ***human computer interface*** or human element is positioned in the decision loop to monitor progress and insure overall reliability of the data fusion system.

Colonel Charles Sadowski (USAF, Retired) is an automatic target recognition expert and research sponsor from Headquarters Air Combat Command, Directorate of Requirements. Colonel Sadowski has worked of many targeting applications for the U.S. Air Force and offered additional advice for constructing a robust hierarchy that would improve upon the weaknesses of current target recognition applications. Colonel Sadowski suggested a strategy that starts with general descriptor classes and then incorporates select details as needed, rather than current approaches of training classifiers based on specific and detailed features and then struggling to generalize them upward.

Pulitzer Prize winner, Douglas Hofstadter [111] proposed a similar procedure building upon established classes or categories. From his book "Godel, Escher and Bach" [111], Hofstadter provided an example, that follows, of an approach that continually added context to an entity, creating an instantiation of the object that is very different from its original form. First, imagine a publication, and then add enough print until it becomes apparent that it is a newspaper, followed the addition of a unique typescript, and title that morphs the entity into the well-known San Francisco Chronicle. With the inclusion of temporal data, the paper becomes the May 18 edition of the Chronicle. Adding relational links, the Chronicle becomes my personal copy of the May 18 edition of the paper. Now depending on the temporal relevance of this data, it can be the May 18 edition of the Chronicle as it was when I first paid for it or the discarded copy days later in the fireplace burning.

Using a combination of the previous philosophies and advice, the initial face recognition methodology that was focused on creating a refined face space, is modified instead to incrementally build upon the steps of segmentation, processing, and identification. These steps utilize commonly utilized information from the spatial dimension of the image, but also integrate and exploit spectral elements assisting throughout the segmentation, processing and identification phases. Figure 99 shows the result of this evolution and the resulting face space hierarchy.

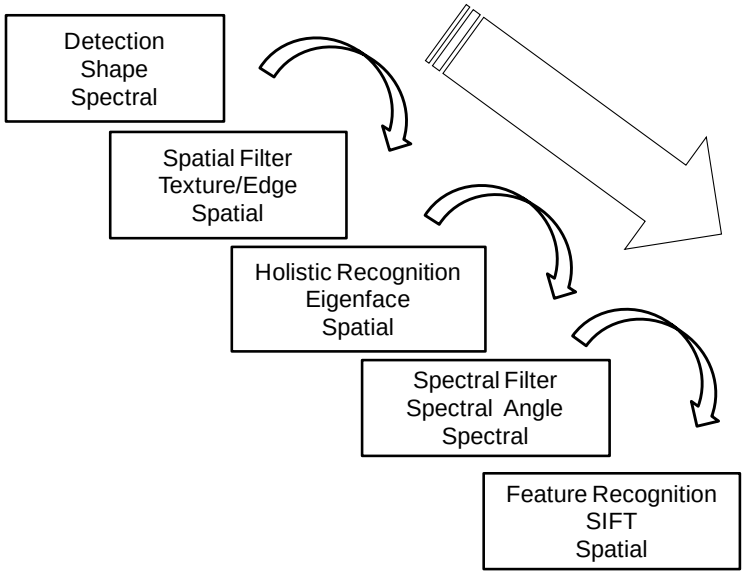

Figure 99: Hyperspectral Face Recognition Hierarchy

At the initial level, a normalized difference index similar to the NDSI is used in the spectral space to locate and segment skin and hair in the face image. The following step of locating or enhancing edges in the spatial space provides textural cues that complement the skin and hair segments. A holistic approach of the eigenface method is applied next to integrate a spatial grayscale representation from the visual spectrum [20].

The ensuing step switches to the spectral domain as spectral angle matching is applied to select areas of the face for matching tissue types. The last step uses Scale Invariant Feature Transform (SIFT) to locate robust interest points from the spatial representation of the face for final identification [60]. The utility of these successive steps were first explored as a progressive filter and later used more effectively in the fusion of matching scores.

Daugman framed the fundamental performance challenge that we faced in creating a face recognition systems [51]. Ideally, these systems and evaluated feature sets should contain the smallest intra-class variance and the largest inter-class variance. Unfortunately, when different face images are captured with a variety of expressions, poses, and illumination, it is common for the intra-class variance to be larger than the relative inter-class variance. With this challenge in mind, it is important to select and incorporate features that remain stable across the multitude of variations that are encountered when utilizing face images.

Using both general and detailed features provides stability for recognizable attributes. The face shapes act in the same manner as they do for the accomplished artist, framing his portrait [5], or the recognizable silhouettes of famous celebrities, using nothing more than the outline from a front or profile view to aid in recognition. At the connecting levels, familiar spatial and spectral characteristics are used in various matching algorithms. Finally on the end of the continuum, the specific SIFT interest points provide scale and rotational invariant details that tend to be tolerant of illumination changes.

$i = 1,\ldots,G$	number of subjects in the gallery
$j = 1,\ldots,G'$	number of probes to identify
$G' \subseteq G$	probe always has a matching face in the gallery
$k = 1,\ldots,C$	number of classifiers or agents
$S_{ijk} = j^{\text{th}}$	score on the i^{th}probe for agent k
$S(i,j) = \sum_{k=1}^{C} b_k S_{ijk}$	total score for weighted sum of fusion of agent scores
If $\min_j S(i,j) = S(i,i)$, $\text{TP}_j = 1$	true positive
$\text{TP}_j = 0$ *o.w.*	false positive
Let $TP = \sum \text{TP}_j / j$	true positive rate

Figure 100: Mathematical Formulation of HSI FR Problem

The notation, shown in Figure 100, can be used to assemble the mathematical framework for the problem addressed by the proposed hierarchy. Continuing with this initial framework the logical next step would be to choose the weightings that maximize the true positive rate. This approach, although novel for hyperspectral imagery, has been replicated for other image types and with other algorithms but has proven limited ability to advance face recognition past its current limitations and challenges discussed earlier. A more adaptive and intelligent framework such as that offered in the QUEST methodology would have to be developed to elevate this initial approach to new possibilities. The means to implement this methodology and the associated performance and efficiency advantages will be investigated through experimentation.

Although the problem is articulated in terms of true and false positive rates, the primary means of evaluation and comparison with published approaches is through the commonly used CMS plot. The CMS plot portrays the true positive rate but additionally

illustrates the attempt in which the correct face was matched. For initial testing, the number of probes will equal the number of candidates in the gallery.

Using the combined expertise cited previously, the original experimental methodology (see Figure 61), was expanded from simply tailoring the comparison space and gallery to an approach that incorporated a incremental strategy that proceeds from general characteristics, that are easily extracted, to more specific and robust features. The next element in the development of this hierarchy is to incorporate contextual cues to construct our version of the May 18th edition of the Chronicle. This instantiation of the face has developed from a simple shape, to a detailed image containing tissue spectral signatures and invariant interest points and soon these features will be linked to provide a contextual representation of a subject's face like it has never been used before.

Application

A review of the basic matlab functions along with supporting detail follows to provide an overview of the underlying algorithms implemented in the hierarchy. The first function, *faceSA.m*, locates the face surface through a combination of a normalized difference ratio and spectral angle. The comparison of resulting face shapes is accomplished with the eigenface algorithm. The next function, *hairNRGB.m*, acts in a similar manner as the algorithm locates hair surfaces through NDVI difference index and has options for refining the detected hair segment using a threshold based on either Mahalanobis distance or standard deviation. The eigenface algorithm compares the

detected hair segment shapes. The outputs of these first two functions are the relative hair and face segments and the relative scores based on Euclidean distance.

Utilizing the previous face segments, the next function called f*aceparameters.m*, calculates common measurements such as area, major axis, minor axis and eccentricity to create a feature vector of measurements. These feature vectors are compared using the Mahalanobis distance between each vector and the mean vector of the gallery. The segmented shapes of the hair and face are subsequently used as a fitted template to extract a grayscale representation of the respective segments. The spatial matching capability of eigenface is implemented with the functions, *facerecognition.m* and *facerecognition2.m*, that compares grayscale images of face, or combined face and hair images respectively. The matching score is a measurement of Euclidean distance.

The spectral elements of the images are evaluated using the following two functions. The first function, *hairSpectrum.m*, conducts spectral matching of hair signatures, but due to the variability of hair samples, is aided with the outlier detection method, BACON, before Mahalanobis distance is measured between subjects. For spectral matching of the face segment, the function, f*acespectralmatching.m*, conducts spectral matching of skin signatures using a spectral angle comparison of the mean signatures of each face. The option to improve results using the BACON algorithm refinement is available for both functions. The output of this matching function is a score representing the measurement of spectral angle.

The interest point matching is accomplished by Lowe's SIFT algorithm [60], [61], in the function, *SIFTrecognition.m*. The function identifies the number of matching interest points, or SIFT keys, between individuals. This algorithm can be applied to the

face image or the combination of hair and face images. The number of matches is normalized and then subtracted from one to represent the best number of matches as the lowest score. This transformation is conducted to produce an output score similar to other functions where a normalized distance closer to zero represents a stronger match.

Assembling these functions together produces a number of comparisons between subjects that parallels the image-processing scheme described earlier by Ando [109]. Proceeding from low level segmenting, through high level processing techniques, details are integrated across different perspectives that form an identification system that utilizes both spatial and spectral information. Using the functions in a serial process, the hyperspectral face recognition (HFR) methodology becomes a functional identification system as illustrated in Figure 101.

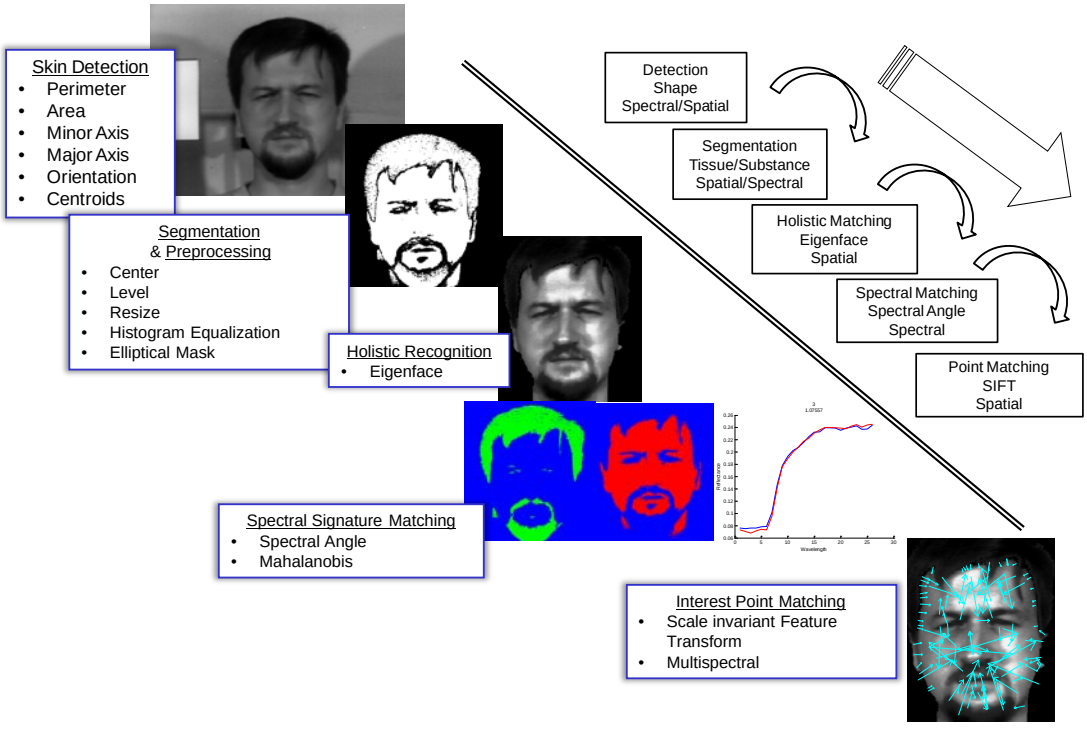

Figure 101: Applied HFR Hierarchy

Applying this methodology using both score and rank fusion illustrates the value of this technique and the results are depicted in Figure 102. At the basic stage, we see the marginal value of shape measurements, displayed as soft features. The accuracy of the algorithms increase as the resolution of the features increase up through the notable SIFT algorithm that matches only invariant interest points. One exception is the surprising utility of the simple method matching the segmented hair shapes. This characterization is helpful for this purpose that assumes an unsuspecting or cooperative subject, but could just as easily be detrimental when a subject is determined to circumvent the system by altering their appearance with a change of hairstyle.

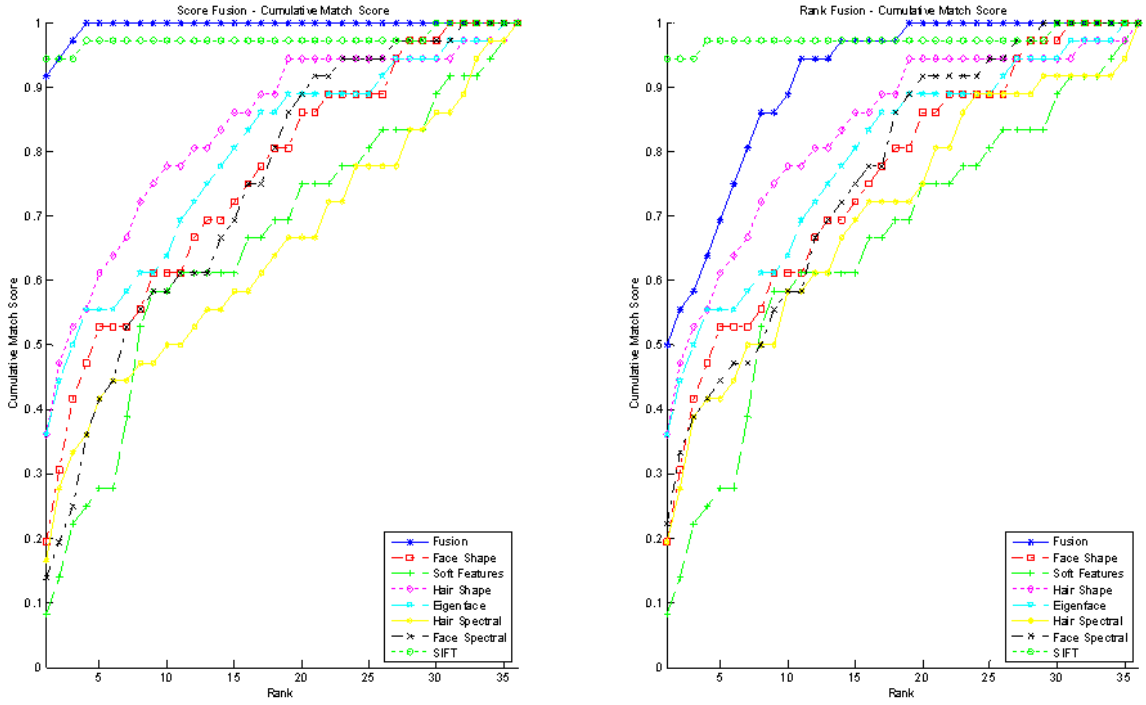

Figure 102: Score and Rank Fusion Results for HFR

To examine the relative contribution of this strategy, a comparison is made with the results obtained by algorithms previously used in hyperspectral face recognition

research. These algorithms include, Pan's spectral face algorithm, Elbakary's spectral matching, and the well-known eigenface method. Dissimilar from their published results, they all encounter difficulty when applied to this challenging data set. On the other hand, the initial results obtained with the serial implementation of the HFR hierarchy shows immediate promise. The benefits of this approach include the removal of human interaction and the accompanying time and effort required for the manual processing of images, as well as the increase in performance as shown in Figure 103. Keep in mind, for this initial attempt, no adjustment or optimization of the fusion weightings where used. This opportunity must be explored in future efforts to pursue maximum performance gains.

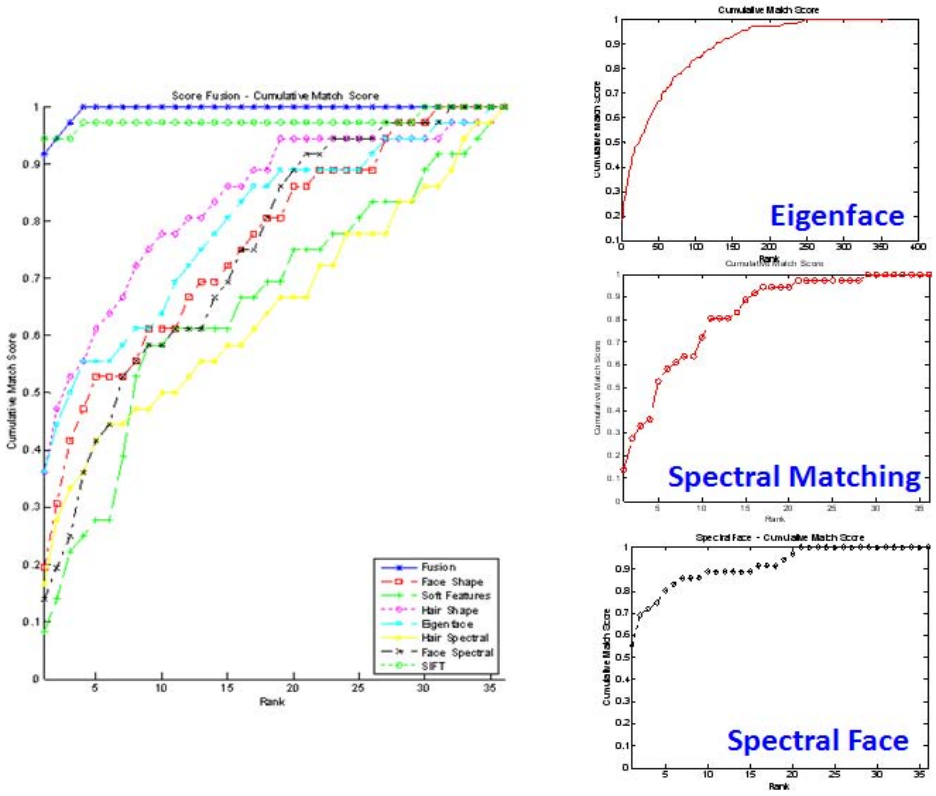

Figure 103: CMS Plot of HFR Methodology vs. Competing Methods Tested on CMU Data

Utilizing the available database to the maximum extent possible, a second group of probes was selected to serve as a validation set. Limited to the subjects that attended at least three separate sessions; the number of validation subjects had to be reduced from 36 to 28 subjects. The actual breakdown is shown below in Figure 104.

16 Subjects have **5 sessions** of data (images with lighting from left 45, center, right 45, total)
- Subject #s: 01, 02,07,14,18,19,20,21,22,23,24,25,30,31,34,36

6 Subjects have **4 sessions** of data (images with lighting from left 45, center, right 45, total)
- Subject #s: 04,05,12,15,26,41

6 Subjects have **3 sessions** of data (images with lighting from left 45, center, right 45, total)
- Subject #s: 08,09,10,11,13,28

8 subjects have **2 sessions** of data (images with lighting from left 45, center, right 45, total)
- Subject #s: 06,16,17,29,33,37,38,40

12 Subjects have **1 session** of data (images with lighting from left 45, center, right 45, total)
- Subject #s: 03,27,32,35,39,42-48

Figure 104: CMU Validation Data Set

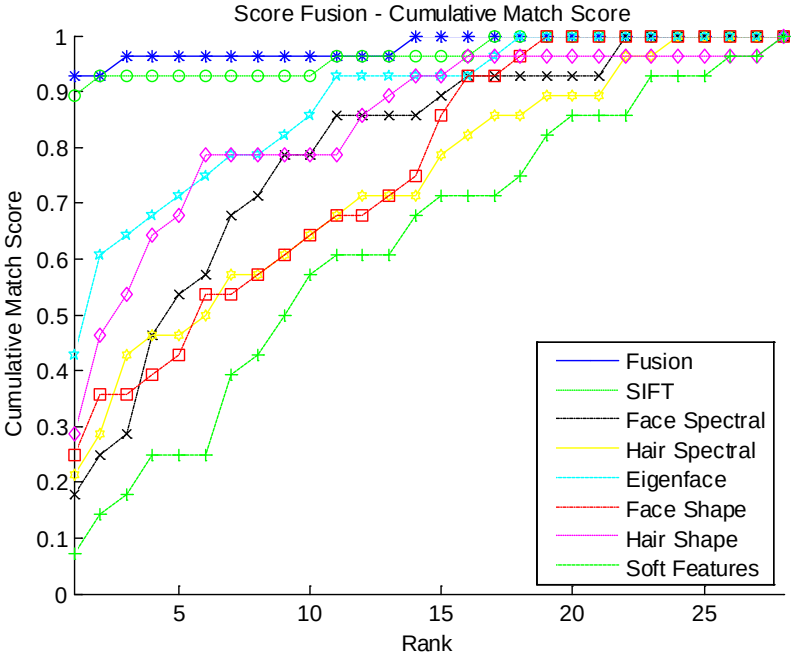

Figure 105: CMU Validation Results

Using the same hierarchy and supporting algorithms the validation confirmed the earlier promise of the methodology, producing very similar results. The significant increase in overall accuracy was maintained over the other leading HSI algorithms. The general performance remained consistent with some slight shuffling of the relative effectiveness among supporting functions. The validation test results are shown in Figure 105. For this limited sampling, it is difficult to make additional observations outside the intended purpose of the validation effort.

Although the increasing capability of modern sensors provides an ability to analyze wider portions of the electromagnetic spectrum, the quantity of data that accompanies this technology brings an associated cost of processing and storage. Operational face recognition systems are often required to process large databases that can quickly diminish the potential value of using hyperspectral data if an efficient approach is not used.

For this reason, several strategies are leveraged to reinforce efficient extraction and analysis of the high dimensional hyperspectral data. The straightforward manner of the normalized index, essentially a ratio of relevant wavelengths, for hair and skin segmentation rapidly selects a small subset of pertinent data for processing through the collection of recognition algorithms. Using the resulting segment shape as a tailored template, as opposed to an arbitrary ellipse or rectangular frame, data compression is easily implemented as the span of hyperspectral frequencies are combined to produce an average representation, represented as a grayscale image, for spatial matching tasks.

Moving from algorithmic to feature set efficiencies, the calculated soft characteristics used can be utilized to filter out images from the gallery that are irrelevant

to our search based on distinctive differences between descriptive characteristics. A simple example of this was shown earlier using a NN application on spectral skin signatures that is reliant on the biological melanin content of an individual's epidermis and stratum corneum. As these soft characteristics are extracted and combined, contextual layers are developed that reduce the number of subject that must undergo a complete comparison and provides information similar to Jain's soft biometrics [1]. Recalling Jain's earlier research on soft biometrics, his proposed follow-on research was to develop an automated biometric system that could extract and use soft biometrics to improve biometric modalities. This research begins to accomplish this task and reinforces the findings of his research.

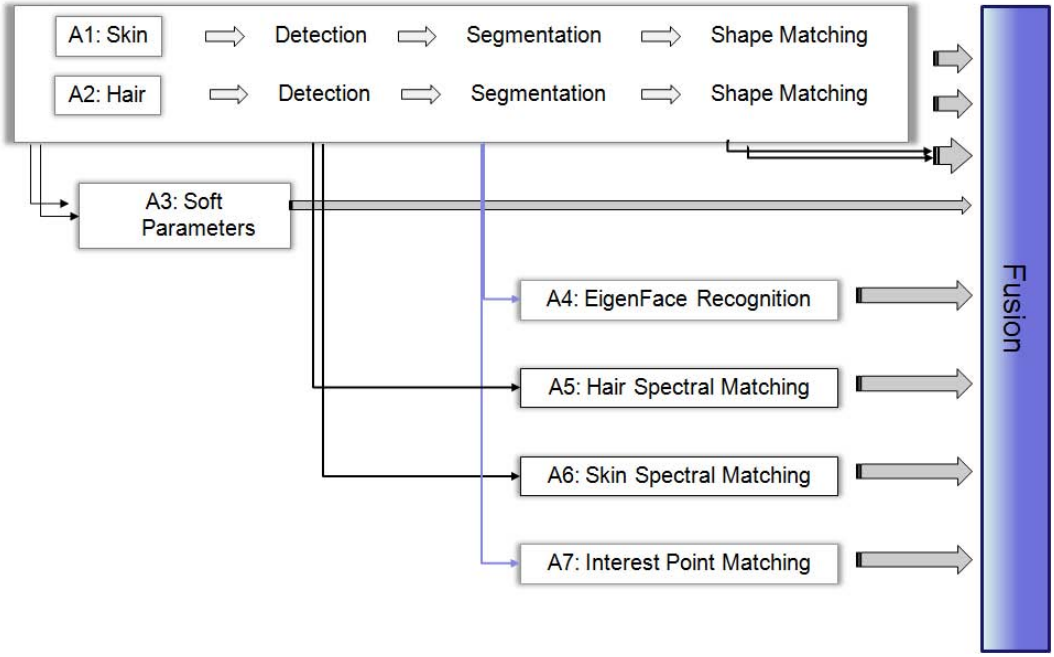

Figure 106: Recognition Agent Interface

Using only a subset of available data, along with some applicable abstractions, links are generated and information is shared, creating a knowledge interface among the respective algorithms. This structure is intended to replicate the QUEST methodology and agent relationships discussed earlier. Referring to Figure 106 and recalling in the QUEST methodology, we attempt to develop a intelligent computational system that captures the advantages of qualia-like representations. These representations provide a context of biometric characteristics, fiducial features, and cues across the electromagnetic spectrum. The supporting functions are linked agents that communicate information that is the context for other agents and in turn, this contextual information can improve performance and efficiency.

In summary, this agent interface uses the normalized difference indexes of skin and hair to create segmented layers that are used singularly and in combination to create information layers and a contextual framework utilizing cues from both the spectral and spatial realms. This representation can be used as a front end to enable automatic preprocessing and integrated with subsequent recognition applications. This information sharing is valuable in aiding performance as well as addressing the weaknesses of circumvention and uniqueness.

Adaptive Feedback

To incorporate the ability to evaluate the most relevant comparisons over time, adaptive feedback loops were added to the established hierarchy and are depicted in Figure 107. The feedback loop for the adaptive gallery is included to examine the impact

of changing the candidate gallery to incorporate only the most likely candidates. This procedure involves reducing the gallery size by removing the lowest scoring subjects. This process is applied only for subject matching scores that fall below a user specified threshold.

A multi-look functionality was added to test additional probe images if they are available. This facet can represent a temporal dimension that would be provided by the multiple probe images that are available through multiple image captures or with hyperspectral video that would obtain a series of face images over time. Both feedback loops can be active or applied individually. Recalling the insight gained by Sinha's research [4], the comparative bias between images over time has important implications in human recognition and may play an important role in computer recognition as well. This new construct complete with adjusting feedback loops is now significantly different from the conventional biometric system illustrated in the beginning of Chapter 3.

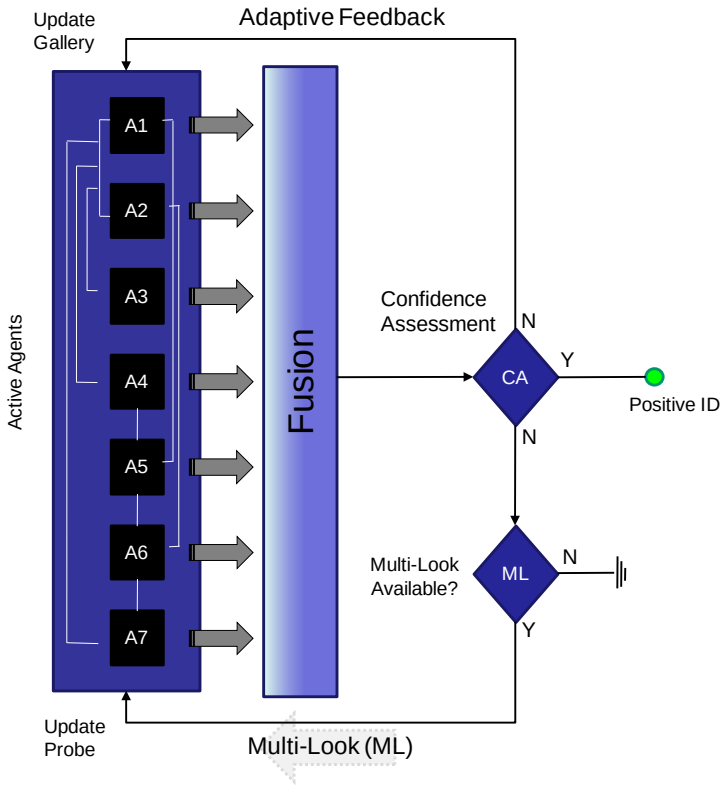

Figure 107: Adaptive Logic

Real World Requirements

At the 2007 IEEE Conference on Biometrics: Theory, Applications, and Systems (BTAS), Dr. Michael C. King spoke as the Director of the Intelligence Technology Innovation Center. Now known as the Intelligence Research Projects Activity (IARPA) Smart Collection Program, this organization is responsible for guiding the technology development of the federal intelligence agencies. Dr. King declared that the development of future biometric systems should concentrate on extending the current capability in three dimensions to make them useful to our nation's intelligence areas. The areas of capability that need improvement are identifying subjects at greater distances, during

subject movement and when subjects are uncooperative and try to circumvent recognition systems. Current biometric capability in this domain, shown in Figure 108, resides near the origin of this axes and additional research is required to extend capability in all dimensions.

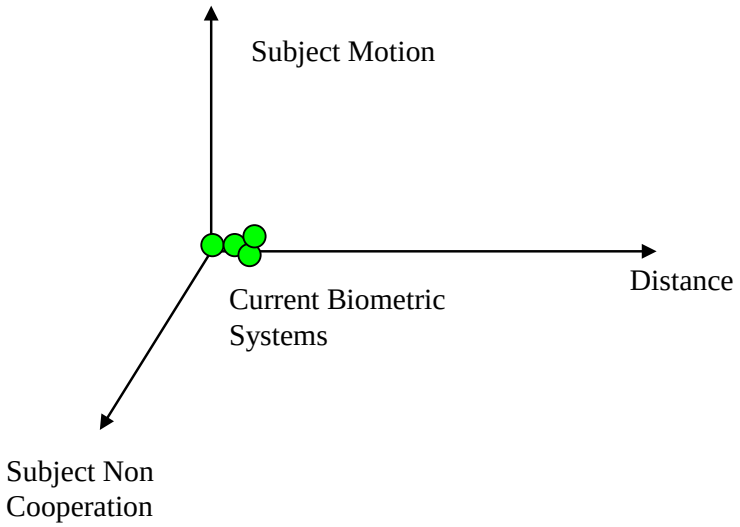

Figure 108: Extension of Biometric Research

This need for face recognition has been aided by this research in the following ways. The use of hyperspectral imagery is a well know technology for remote sensing applications but has not been operationally applied to face recognition. Subject tracking is already possible through the use of IR imagery but the inclusion of information from other wavelengths can improve recognition capability of IR based applications. This potential has already been shown with the work of Bowyer [89] using grayscale images in conjunction with IR imagery. Similarly, motion and tracking capability can be enhanced with a hyperspectral data and the proven ability to detect and segment human skin and hair.

Through the use of soft biometrics and the QUEST hierarchy, distance recognition can also be enhanced. Earlier detection and screening potential made possible through the broad use of the electromagnetic spectrum can provide useful cues before adequate resolution is obtained for positive identification. This capability can also enable the classification of uncooperative individuals at a distance through the use of spectral information from a number of surfaces and perspectives. The ability to evaluate the spectral signatures of individuals and substances allows the detection of spoofing attempts to include disguise or obscuration.

Following the incorporation of selected soft characteristics, additional features are used to alleviate many of the aforementioned challenges of circumvention and uniqueness. With the size and shape of the skin and hair segments already identified, spectral signatures are automatically collected from these segments, processed and used in various matching algorithms. Previously, a series of cumulative match curves for the different tissue was shown (Figure 76) and although these applications do not conclusively identify an individual with this data, it provides an additional reduction in uncertainty, as identities are cumulatively determined. Additionally, the unique signature of the hair, compared against human skin, can help locate and segment the important portions of the face including hairlines, eyebrows, beards and mustaches. At the same time, the unique spectral characteristics can also serve to highlight inconsistencies such as individuals trying to alter their appearance with hair extensions, dyes, or wigs. The same is true of an individual attempting to circumvent a recognition system by using makeup or prosthetics to disguise their appearance.

This performance capability addresses the challenges of non-cooperation and other pressing needs as outlined by our nation's intelligence agencies. The ability that Dr. King is seeking is contained in the potential of HSI and the framework outlined by the QUEST methodology. Through this methodology the capability operational envelope for face recognition can be extended beyond traditional applications in all three of the desired areas.

The *Handbook of Face Recognition* [105] summarizes the implementation of face tracking and recognition from video data. Most systems primarily focus on face detection and tracking and subsequently implement recognition after an image meets certain size and pose requirements. In this multi-look implementation, the considerations of size and pose are minimized with the use of the scale invariant feature transform that identifies both whole objects and partial segments at various orientations and scales. The incorporation of both spatial and spectral classifiers supplements this robust functionality.

A common critique of "still to still" recognition approaches that the attempt to solve the tracking and recognition problems sequentially is that they do not exploit the temporal aspect of the video data. The design approach in this research solves the tracking and recognition problems simultaneously while using temporal information through the use of hyperspectral data and the feedback processes that adjusts gallery size and composition, injects new and subsequent images, and offers the possibility to use trend information and decision thresholds for an intelligent and adaptive system. The incorporation of the unified probabilistic framework by Chellappa [105], applicable to any still to still application, should be a considered for future enhancements and analysis.

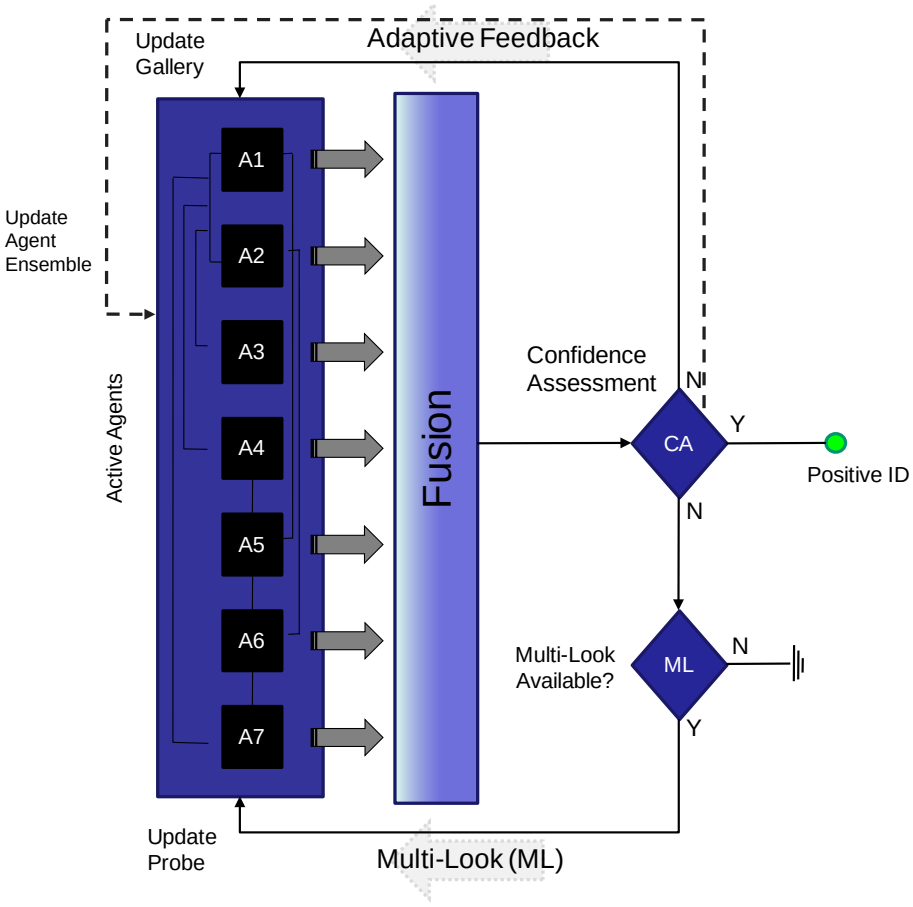

Figure 109: Ensemble Enhancement for Future Research

The final adaptation incorporated into this methodology are control variables to enable the selection and weighting of agents used in the fusion process. Research by Chawla [33], [34] and Kuncheva [86] have highlighted the importance of randomness and diversity in the creation of classifier ensembles. The systematic and random selection of these active agents should be a continued area of research. The appropriate adjustment to this framework would be the addition of a third feedback loop for updating the agent ensemble selected as pictured in Figure 109.

Graphical User Interface (GUI)

Figure 110: HSI FR GUI for Data Exploration and Strategy Development

To facilitate data analysis and assist with visual interpretation of the results, a Matlab-based GUI tool was designed to operate and test the designed facial recognition system. The GUI, pictured in Figure 110, is a direct parallel to the architecture presented earlier. A user can select the active agents, enable desired feedback loops, and designate fusion weighting schemes while simultaneously analyzing results. The GUI displays the probe image to be matched as well as the best current match directly opposite it. Below these images is a line-up of the top ten matches displayed in a thumbnail depiction along with their true identity and associated matching score.

Viewing the results from each individual agent is permitted by selecting the algorithm of interest in the "Results to Display" drop down menu. If feedback loops are employed, the user can view the results from the specific iteration through the "Gallery Set Results to View" menu. The pictorial results are viewable in either grayscale or color images.

To review the quantitative results, the user can choose from an assortment of summary products. These products include cumulative match score plots, box plots or histograms of the relative local scores for a single probe or a summary view of all matching scores for the entire test set.

For computational planning purposes, Matlab's multiple processor pooling was employed on a dual quad core computer with 16GB of RAM. The processing requirements of the hyperspectral data along with the possible array of algorithms benefited from the occasional use of the parallel processing capability of this system and the built in functionality of Matlab's parallel processing toolbox. To review results on any computer, a utility tool was developed to enable users to view saved results of any prior run by loading a single file. This utility file will display the active agents used, types of feedback loops employed, the associated weighting schemes, along with the ability to view all related results and images. The user of the utility tool is notified if the current computer can support the computational requirements of the recognition software suite. A sampling of results will be reviewed in the next section to demonstrate the capability of the QUEST HFR methodology and the functionality of the GUI.

V. Results and Findings

Starting with the initial round of results that utilized a weighting that was biased towards the most effective SIFT algorithms and a combination of both hair and skin segments, the CMS reveals the ability to correctly identify all candidates in the test set. This performance alone exceeds previously devised strategies as applied to the CMU dataset that resulted in the following correct matching capability: Spectral Face (Pan) 56%, Eigenface (Pentland & Turk) 20%, and Spectral Matching (Elbakary) 14%, as depicted earlier in Figure 88. This performance is achieved despite the difficulty with data quality, appearance variations among the subjects during separate sessions, movement and changes in facial expression during image capture and the alteration of hairstyles and the addition of sunglasses. Figure 111 depicts the overall results.

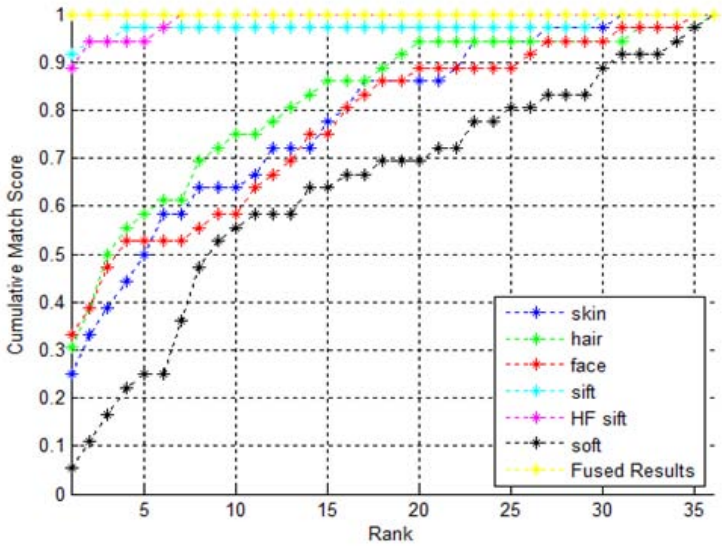

Figure 111: Cumulative Match Score Results Including Score Fusion

Although the performance is encouraging, there is insight offered by evaluating the quality of the match. Additional data analysis functionality was incorporated into the user interface for this purpose.

A combination of score and rank fusion strategies were tested with the most effective being a weighted score fusion strategy. The performance depicted in the previous cumulative match score was also aided by the interaction of contributing sub-methods. The ability to process and match both the independent segments of hair and face and combined images through the spatial and spectral domains provides an added performance boost over earlier strategies.

The additional functionality of the adaptive gallery and the multi-look feedback allows the recognition process to continue until the results meet an acceptable threshold or confidence level. During each repetition, the gallery size is reduced by eliminating the poorest gallery candidates and implementing additional probe images at each iteration. Through this repetitive process, classifications with the poorest matching scores are re-evaluated to confirm the correct identification. The recognition capability that results provides matches that are not only correct but also obtained with an acceptable level of confidence. This process is investigated further with the help of the interactive graphical user interface.

Using the *boxplot* function of matlab, the matching scores can be quickly explored without making unjustifiable assumptions and the quality of the matches can be inspected and intuitively characterized. The box plot representation was chosen because of the ease in interpretability and its robustness to various distributions and outliers. With this single display, only a quick inspection is needed to visually evaluate the quality

of the match as well as the spread of gallery scores associated with the selected probe. With the side by side arrangement, the relative quality across the entire set of probes and the global population of comparison scores is obtainable visually. This ability will be important as assorted links and feedback loops are experimentally tested in the search for performance enhancements. A quick review of this method first advanced by McGill, Tukey and Larsen follows [114].

The median is annotated as the middle red line, the box marks the 25th and 75th percentiles, while the whiskers mark the boundary of the data points, not including any points identified as outliers. The determination of outliers utilizes the *boxplot* matlab function and annotates the points identified as outliers as red plus signs (+). Using the 25^{th} quantile (q_1), 75^{th} quantile (q_2), and the whisker length (w), set to the default value of 1.5, outliers are designated such if their score is larger than $[q_2 + w(q_2 - q_1)]$ or smaller than $[q_2 - w(q_2 - q_1)]$ which equates to approximately +/- 2.7σ [115].

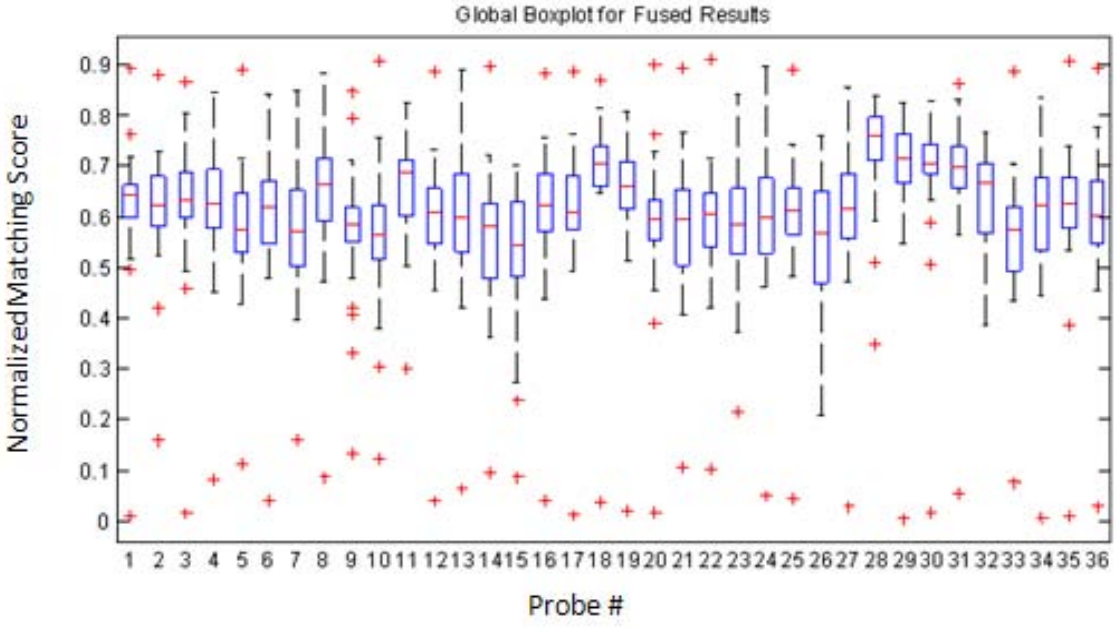

Figure 112: Box Plot of Match Scores for All Probes

With this functionality, the span of all matching scores can be reviewed for quality and anomalies. Figure 112 depicts the entire set of all matching scores, delineated by probe to display the range of and quality of matches for the entire test set. Individual box plots are displayed on the GUI to the right of the closest match for the particular probe selected as shown in Figure 113. In this case, the evaluation of probe number 33 shows a typical strong positive match with a score obtained between zero and 0.1 which is both distant from the median and 25th quantile and as such, considered as the only low valued outlier by the box plot function.

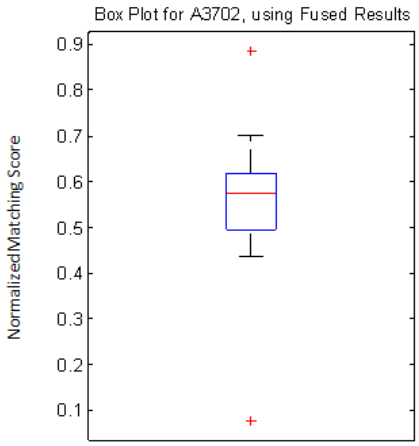

Figure 113: Box Plot of Match Scores for Single Probe

An inspection of the images located in the "Top Gallery Matches" section below the probe and best match, shows the top ten matches and their respective scores. A quick review of this information offers the user an indication of how well the system is operating. The illustration in Figure 114 shows a depiction of the subsequent ranked matches combining the image, the image identity label (see earlier section on CMU data for explanation) and the respective score.

Figure 114: Visual Depiction of Match Candidates

For this particular probe, the top 10 matches happen to possess similar characteristics with short hair, no glasses, similar pigmentation and with the top nine being male subjects. The intent behind some of the selected features is to integrate soft biometric qualities into the hierarchy. In this particular case, the general descriptions appear to be consistent in snapshot of the initial round of results.

Although the box plot and visual depiction of the matches are valuable, further evaluation may require a more descriptive representation of the matching scores. With the creation of a histogram, both the distribution and separation from the mean and standard deviation can be evaluated. This option is available through the selection of the appropriate toggle button. The global distribution toggle button displays all scores from the data set, where as the local distribution toggle button display a histogram for only the scores associated with the selected probe.

The distribution for all match scores in the entire test set and the distribution of the local scores for the current probe are shown in Figure 115. Much like the box plot, the histogram is useful in illustrating the separation between the scores and outlying scores on both extremes. Unlike the box plot, the histogram depicts the relative allocation of scores along the scale.

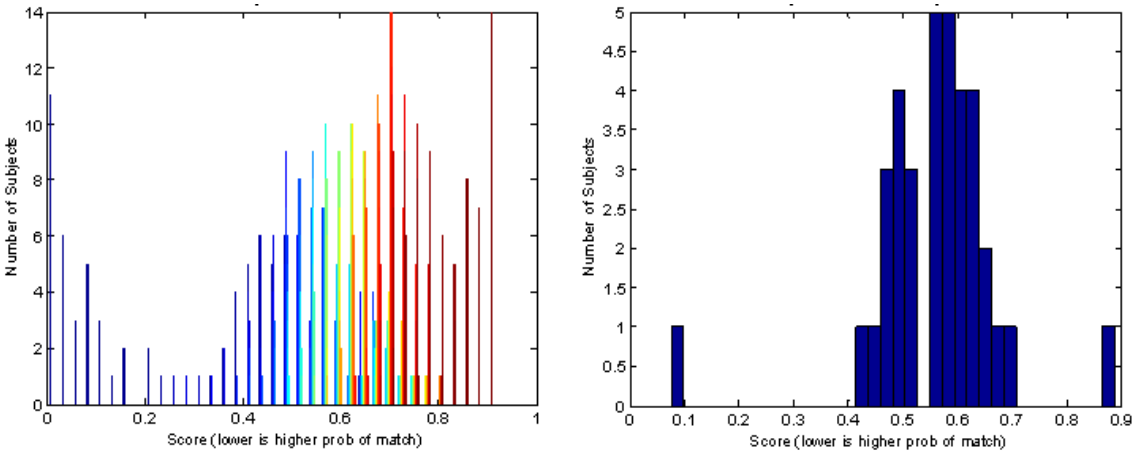

Figure 115: Distribution of Match Scores (Global and Local)

To explain the functionality of feedback loops of the hierarchy as they relate to the GUI, an example of the iterative processing of one of the matches will be reviewed. During the early stages of this research, a goal was set to create a tailored comparison space for improved identification. The logic behind this strategy assumed that shape comparisons for hair and face, as well as the various eigenface comparisons should adjust as the gallery of candidates is tailored. The mean image and subsequent basis set will gradually alter forcing changes in the comparison scores and rankings.

This adaptation can have a significant effect on the underlying characteristics collected by the agents and ultimately the overall results. The comparative matching scores are relative and dependent on the gallery make-up and the extracted descriptive features. The changing gallery leads to an adaptive feature set and an increasingly relevant context for our comparisons. The shrinking subspace and adaptive comparisons produce the tuned face space sought after in early stages of the research as shown in Figure 61.

Figure 116: GUI Evaluation of Sub-standard Matches (Subject 12)

An investigation into one of the higher match scores will provide an example of the effects of the adaptive processes. By inspecting the box plot of all probes (Figure 112), the resulting scores for subject number 12 (11th probe), is noticeable because of the value of the best matching score. Using the GUI, Figure 116, for closer inspection of the initial match for this subject the underlying results will be inspected more closely.

Although the correct match has been made, the score of 0.301 is higher than typical matches, greater than two standard deviations above the mean of 0.092, for the other probes in the test set. There is only one other score of worse quality (Subject 36 or the 32nd probe) using this same criteria and this probe will be investigated later using two separate strategies. By inspecting the box plot, the score is identified as the only outlier on the lower end of the score range. The nature of the match does not seem troubling as the top candidates are visually inspected, revealing an array of male candidates, with one exception, and all with relatively similar characteristics including short hair. A closer

inspection of the score distribution, Figure 117, does not indicate any problems other than a closer proximity of the best matching score to the remainder of the candidate scores.

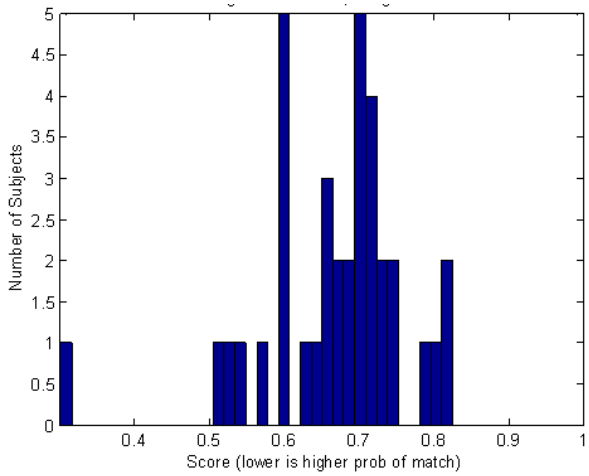

Figure 117: Distribution of Substandard Match (Subject 12)

Transitioning to the next iteration, a reduction in the gallery is made by selecting the toggle button for the Gallery Set Results to the first reduction. This iteration only removes two of the candidates with the largest matching scores. By inspection of the GUI (Figure 118), the resulting box plot and histogram (Figure 119), little has changed except for the removal of the high end scores, which are undetectable by the positioning of the box plot whisker at the far edge of the match score range.

Figure 118: GUI Display after Reduction Step - Negligible Improvement for Subject 12

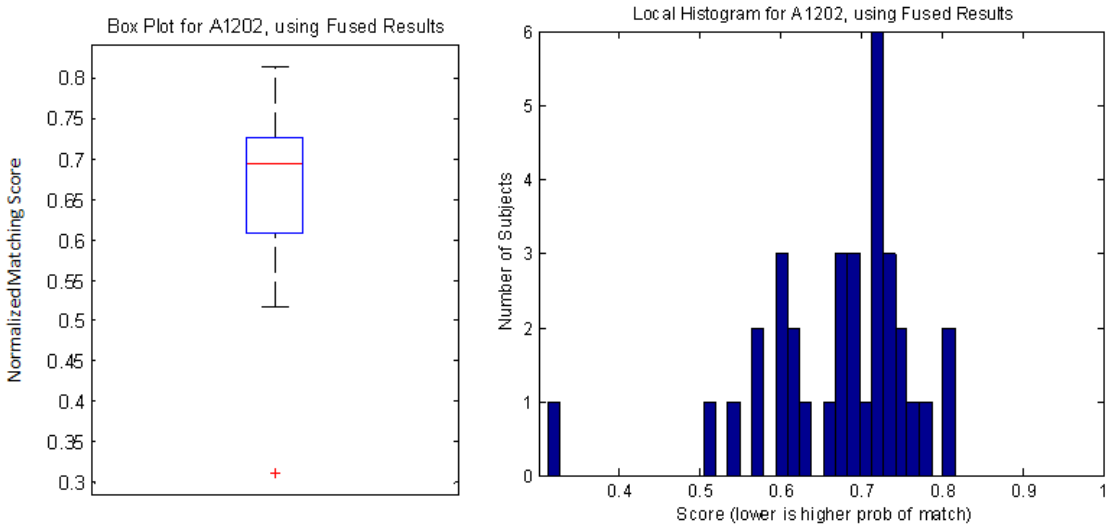

Figure 119: Box Plot and Distribution of Matching Scores - For First Reduction for Subject 12

Looking at the global box plot (Figure 120), shows the relative comparisons between probes and the likely candidates for an additional round of processing. Notice that all but one of the identified outliers on the high range of matching scores have been removed. Our subject of interest, is depicted as the 6th probe below.

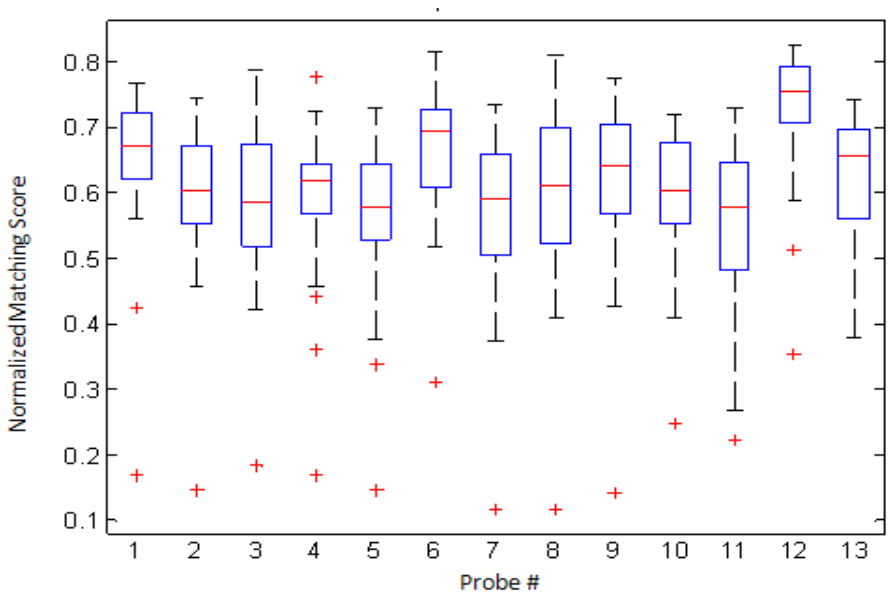

Figure 120: Box Plot for Evaluation of Reduction Set

For this next round, the respective gallery is reduced once again and a new probe image is used if one is available. For this round of matching, the results are more encouraging. In addition to the noticeable change in the score, now 0.097, there is a more significant separation (4.26σ) between the correct match and the other gallery scores as shown in the both the local box plot and local histogram (Figure 122).

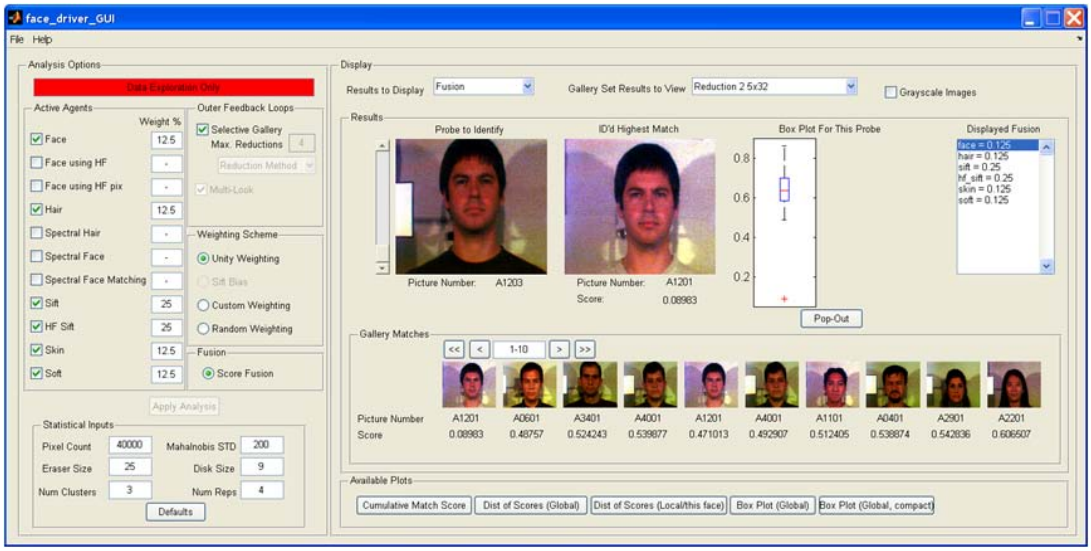

Figure 121: GUI Display for Second Reduction - Significant Improvement for Subject 12

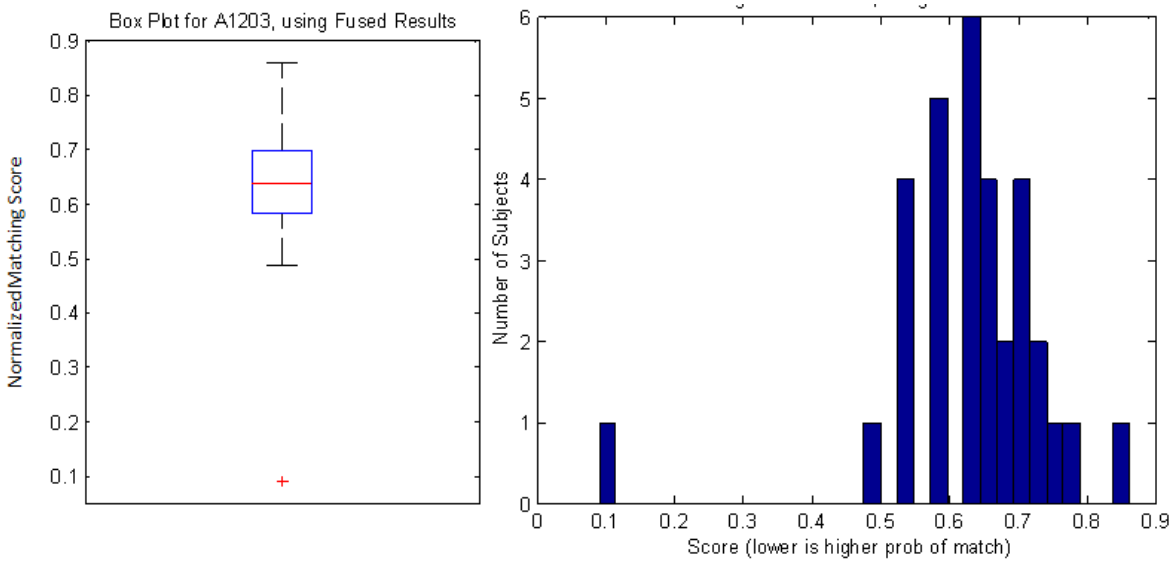

Figure 122: Box Plot and Distribution for Second Reduction (Subject 12)

Evaluating the quality of scores across the remaining test set using the global box plot, displayed in Figure 123, reveals the most likely candidates for additional iterations of this process. Subjects 1 and 2 appear to be both a satisfactory score, less than 0.2, and also with a separation that identifies them as low value outliers compared to the alternative matches in their respective gallery. A point of clarification here is that each probe's gallery has changed independently of the others based on their respective scores. So at this point in time the we are looking at potentially 5 different galleries, albeit a majority of their member would be in common since they started from the same original set of subjects. The remaining subjects, 3 through 5, could be handled either as accurate match or a candidate for further processing depending on the chosen logic and threshold of the user.

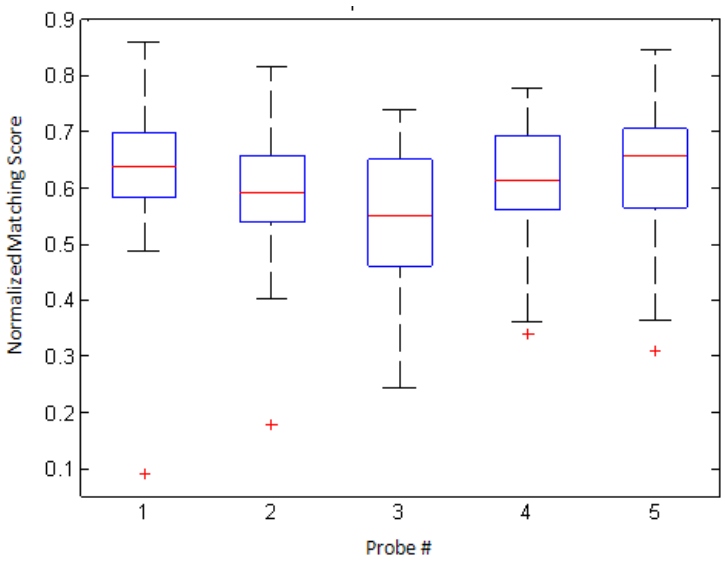

Figure 123: Box Plot for Final Reduction Evaluations

For actual implementation of this system, the previous process would be allowed to run to completion where all matches would achieve a satisfactory level of confidence

is achieved or be categorized as a non-declaration (NDEC) candidate. The purpose of this example was not to determine a specific confidence level or metric but instead to illustrate the value of this multi-dimensional, contextual, and iterative approach for face recognition.

This process was applied across many combinations during the course of this research effort with numerous more on the drawing board. Results from several variations of the combination just reviewed are shown below in Figure 124. As depicted, the results of all these combinations and simple weighting schemes is encouraging with only one subject being misidentified in just the first round of applying the QUEST HFR methodology.

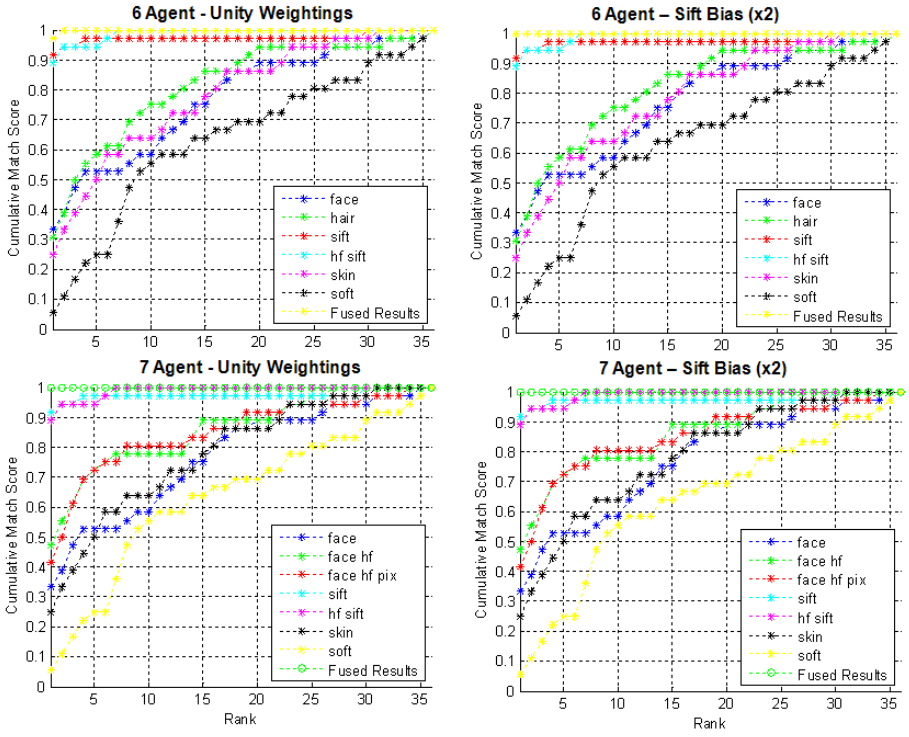

Figure 124: CMS for Various Fusion Combinations and Weightings

To follow-up on the one subject misidentified in all of these test rounds, the quality of the match will be reviewed further to see if this match score would be singled out for additional processing and how the identification score changes through subsequent iterations. At first glance, the matching score is neither an outlier or among the top half of the best matches based on the overall matching score (Figure 125). In fact, based on other best match scores, the matching score is 2.59σ above the mean of the top matches (0.112) and a likely candidate for iteration by most standards.

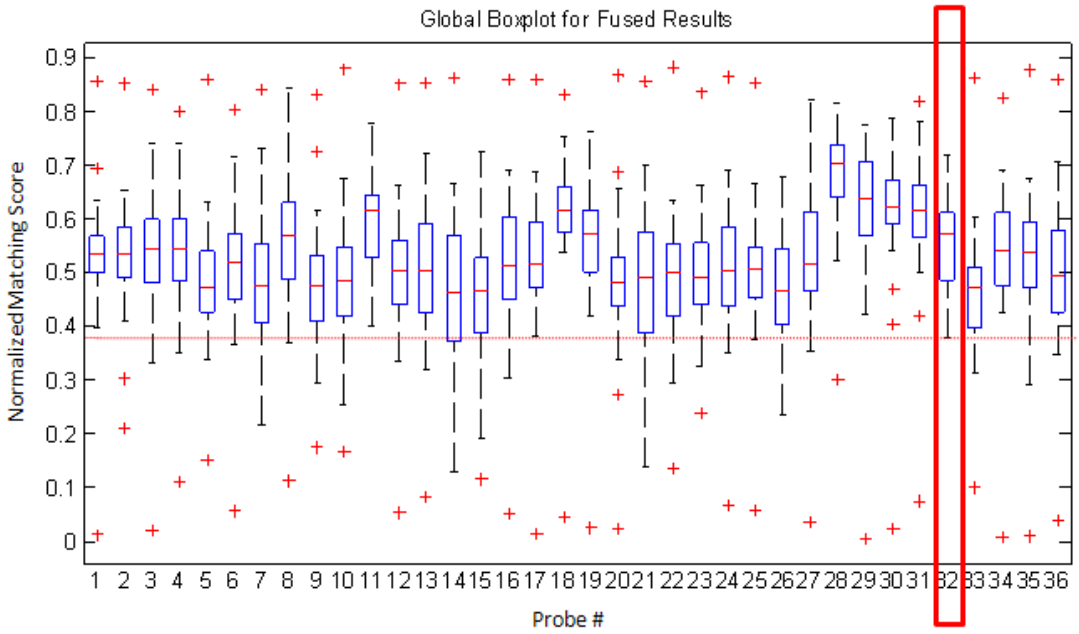

Figure 125: Global Box Plot of All Match Scores for 6 Agent Unity Weighting

The distribution of scores are scattered along the normalized scoring scale with virtually no separation observable. As the methodology is applied, the distribution of scores is shown in Figure 126. Moving from the initial round, depicted in the upper left corner, the scores remain stable in the second round (upper right), start to migrate with the third round (lower left) and finally separate in the fourth round of processing. The

lowest matching score is clearly removed from remaining cluster of scores and upon further inspection we see that the correct match has now been made (this was initially misclassified in the first round) and at a level 3.11σ below the mean matching score.

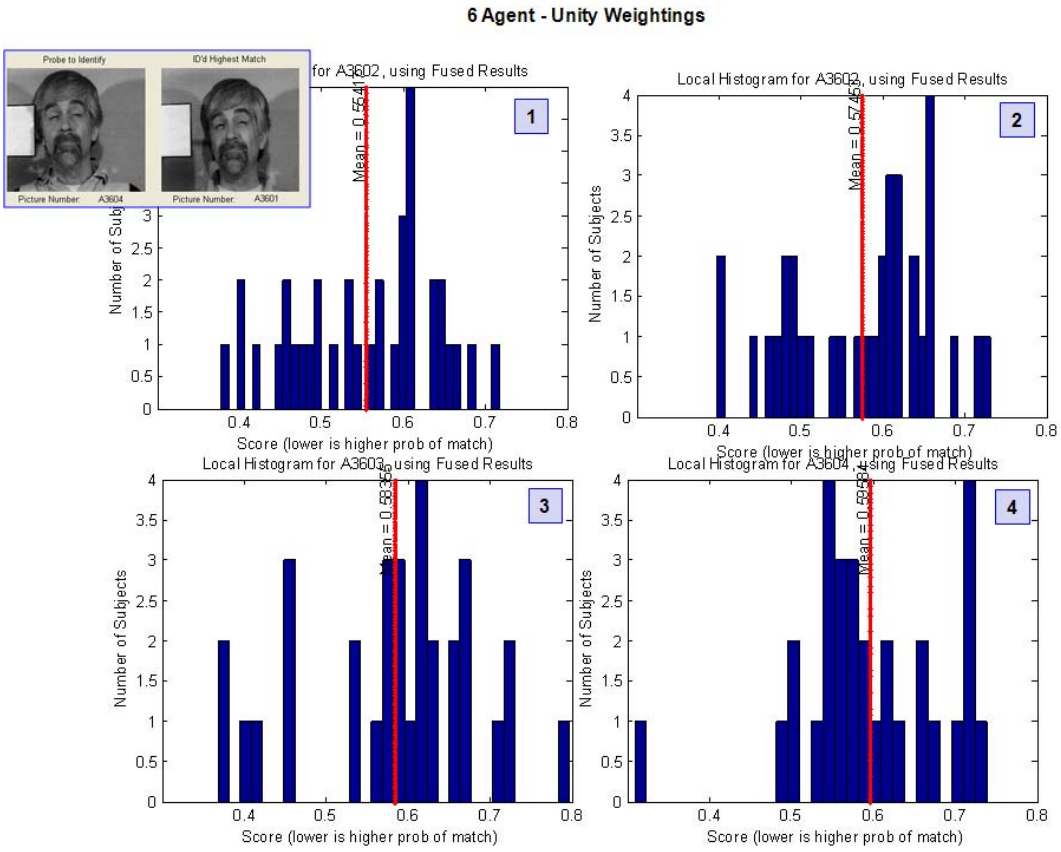

Figure 126: Progression of Score Distributions for Subject 36 (6 Agent/Unity Weighting)

To illustrate that this is not an anomaly for this agent combination (six agents and unity weighting), another inspection of this methodology is made for the algorithm combinations utilizing seven agents and SIFT biased weighting. The SIFT biased weighting strategy leverages the utility of the SIFT algorithms with and increased weighting of two and half times the other agents. The matching score results after the

first round of matching are shown in Figure 127. Again the same subject, number 32 in the figure, is among several likely candidates for a closer inspection.

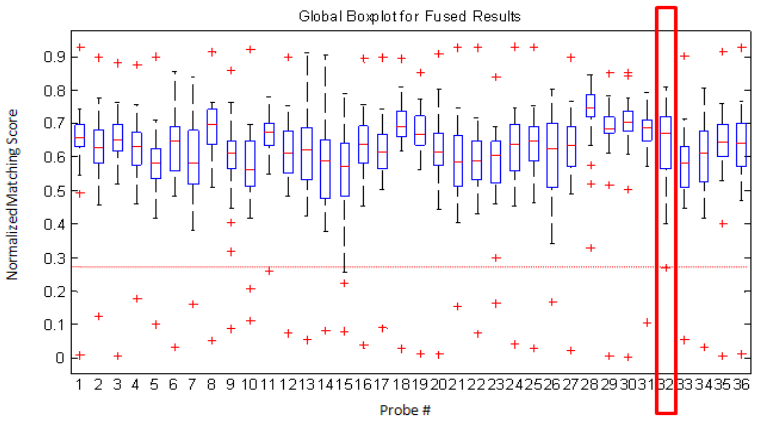

Figure 127: Global Box Plot of All Match Scores for 7 Agent SIFT Biased Weighting

Following the same routine, that utilizes both the adaptive gallery and multi-look functionality, we see the iterative grouping of the matching scores and the separation of the correct match in the fourth round with a distinctive matching quality.

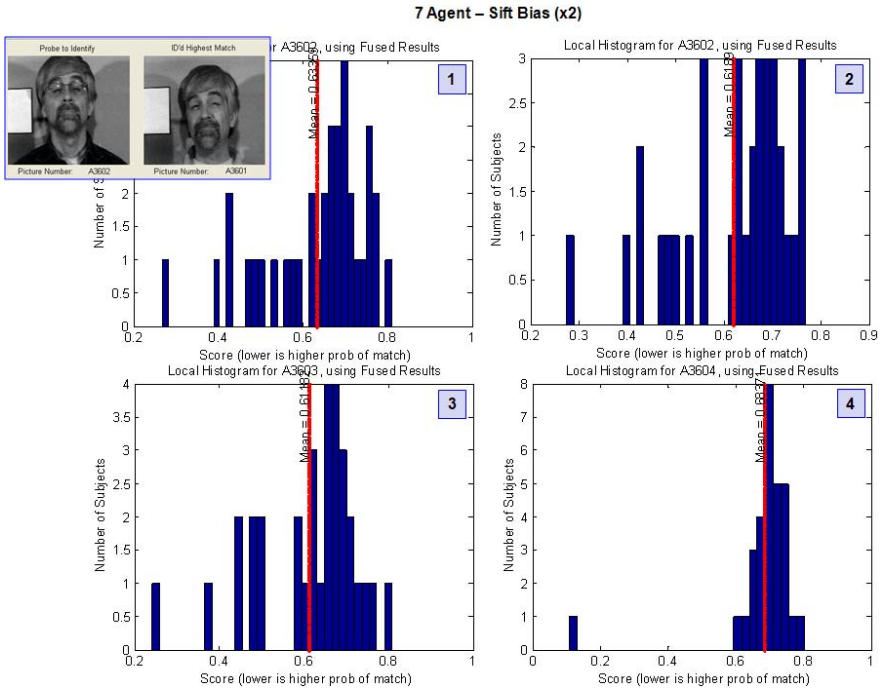

Figure 128: Progression of Score Distributions for Subject 36 (7 Agent/SIFT Biased Weighting)

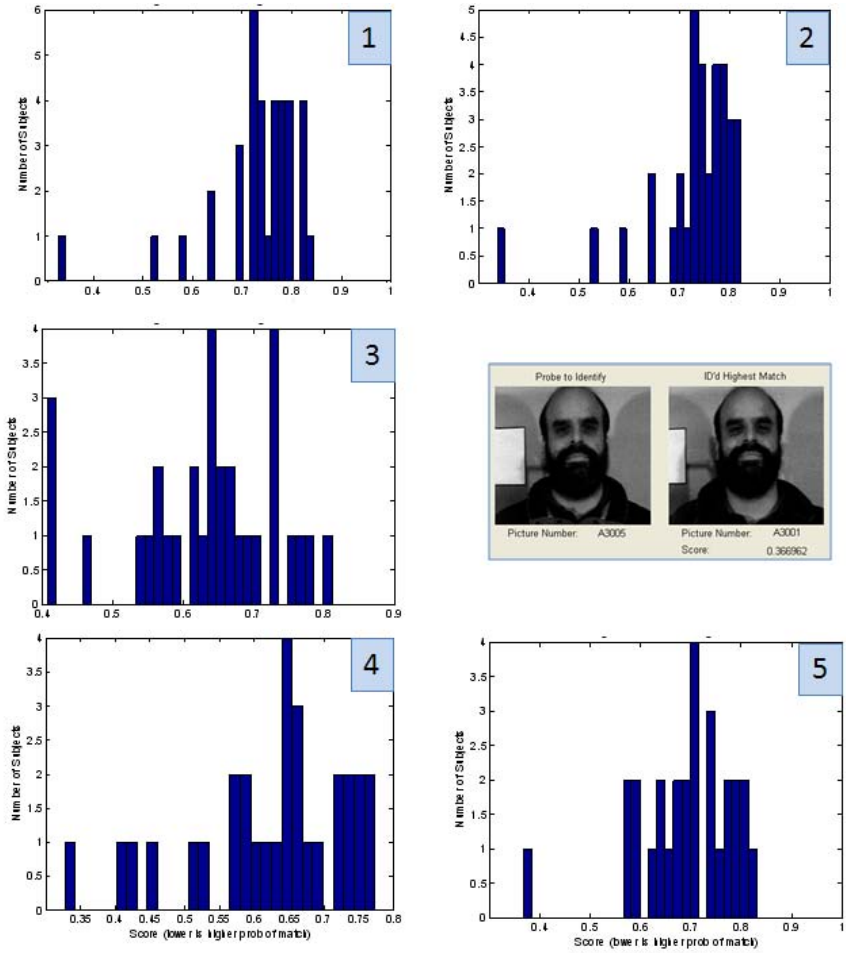

Figure 129: Recognition Progression for Subject 30 (7 Agent SIFT Biased)

In an effort to be exhaustive, a final example using the worst match score (>3σ) from all the testing accomplished is illustrated in Figure 129. The progression in this illustration works through the complete span of images available in the CMU database, in this case 5 rounds, and tests the iteration limit of this database. An interesting and expected result not shown in this depiction is the shuffling of the next closest matches as the face space is tuned through this adaptive process.

VI. Conclusion

The field of face recognition continues to be a highly researched area resulting in the development of many capable techniques and methods. As the use of these systems grow, so too will the number of attempts to circumvent these tools and avoid detection. Countermeasures to circumvention usually take the form of multimodal biometric systems. Unlike many algorithms found in the literature, operational recognition systems are required to operate in uncontrolled environments challenged by various levels of cooperation from the intended subjects. The development of this methodology considers these challenges as it seeks to exploit the electromagnetic spectrum beyond the commonly used visual and IR imagery.

Hyperspectral imagery, already widely used for remote sensing, can be a technology that spans the requirements for accuracy and robustness for environmental variations and circumvention attempts. If hyperspectral technology is going to provide a solution to this challenges an overall hierarchy and processing methodology must be devised to efficiently parse large amounts of collected data. This will become more evident as hyperspectral sensors expand to video imagery and pattern recognition research advances to intelligently process and analyze this data.

In this research effort, the goal was to develop a hierarchal approach to efficiently process hyperspectral face data and demonstrate the benefits of integrating the spatial and spectral domains of imagery. The preliminary fusion results verify the value of

incorporating soft characteristics obtainable into face recognition methods. Even with the undeniable uniqueness of every human being, no single metric or feature has demonstrated the non-intrusive ability to identify all individuals in uncontrolled environments across large populations using a single modality. A demonstrated alternative to this approach may be to fuse contextual or complimentary information in an efficient architecture that enhances effectiveness. This general process is captured in a face recognition biometric system that offers performance and robustness and at the same time mitigates common weaknesses of face recognition applications.

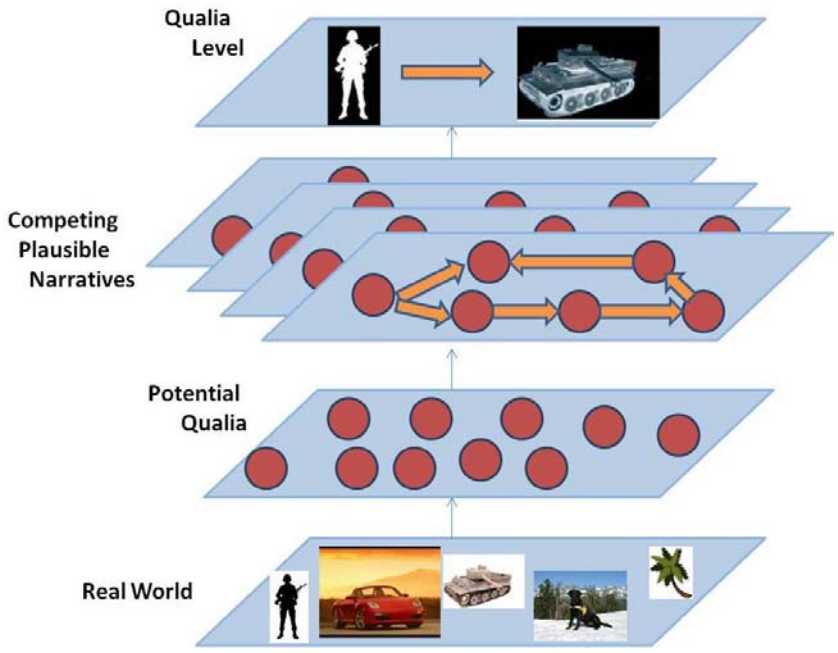

Figure 130: QUEST Situational Awareness [120]

The QUEST methodology and the various levels of awareness (Figure 130) was implemented with many of the key characteristics such as fusion, feedback, qualia, context, general to specific, hierarchical architecture and time incorporated into this

approach for detecting, distinguishing and characterizing faces but with a few simple adjustments almost any entity in the environment. This is reminiscent of the JDL fusion framework where three levels of processing progress from locating and tracking entities, to assessing relationships between objects, and finally making inferences about the current situation [104].

A summary of the findings from this research effort and their contributions to the field of pattern recognition, specifically to biometrics and the face recognition. Recalling Jain's earlier research on soft biometrics, he proposed follow-on research to develop an automated biometric system that could extract and use soft biometrics to improve biometric modalities. This research accomplishes the task and reinforces the findings of his research. Where as many leading face recognition systems rely on human interaction, many times significant time and effort, this system utilizes an approach that enables automatic and efficient segmentation of many of the textures and surfaces unlike approaches used commonly today in remote sensing applications. Using this as a stepping off point allows the implementation of our best and most robust matching algorithms to be used leveraging complimentary information that forms the important context for all recognition to be performed.

To address many of the weaknesses of current face recognition applications, the first being uniqueness, is addressed through the simultaneous use of spatial , spectral, holistic, and local features that combine to portray an individuals identity in a robust and invariant representation. It is only with these various perspectives and dimensions that provide the robust capability when faced with noisy imagery, changes in scaling, translation, and rotation or vulnerable to changes in background environment. The next

and probably most exploitable weakness is that of circumvention. The unique spectral signatures of human tissue types to include skin and hair are detectable, recognizable, and distinguishable from make-up, wigs and prosthetics as shown by this research and as well as earlier efforts most notable Pavlidis [54]. Finally, by combining these attributes in a QUEST motivated hierarchy, we are able to achieve a performance advantage in both time and accuracy not seen in hyperspectral face recognition to date. The accomplishments just mentioned come at a fortuitous time as real world events highlight the weakness of current technology and the requirements for future security systems. This research points towards a solution for the same challenges facing our nation's security and intelligence organizations and that of the international community at large. The continued advancement of technology and some of the principles illustrated in this work can help mitigate the increasing threats and challenges of securing national boarders and populations.

A summary of the contributions of this research are listed below.

- Development of a hierarchal approach to efficiently process data (demonstrated using hyperspectral face data)
- Establishment of automatic preprocessing method for hyperspectral data (demonstrated using face imagery)
- Demonstration of the benefits of integrating spatial and spectral domains of imagery
- Automatic extraction and integration of novel soft features (biometric)

- Verification of the value of incorporating soft characteristics (demonstrated using face recognition)
- Establishment of QUEST methodology in face recognition demonstrating the engineering advantage in both performance and efficiency compared to leading and classical face recognition techniques
- Development of interactive environment for the testing and expansion of the framework advanced in this initial research effort

On October 29, 2010, a young Asian passenger boarded a flight from Hong Kong to Canada wearing a mask that transformed him from his youthful identity into an elderly Caucasian and smoothly through the unsuspecting customs and security officials (see Figure 131)[125]. Sometime later aboard the in-flight passenger aircraft, our traveler removed his disguise, where nearby observant passengers noticed the sudden change in identity and fortunately alerted security officials.

Figure 131: Disguise Raises Security Concerns [125]

In reaction, Janet Napolitano, the U.S. Secretary of Homeland Defense [125], acknowledged that prospective terrorists could implement this same tactic. Jamie Smith, a former CIA officer, offered that this created the opportunity for terrorists to move from one country to another undetected [125]." The use of a commercially available silicon mask, by a clever passenger seeking asylum, was able to bypass the cumulative security measures and monumental investments of the last decade. It is a strong possibility that groups besides freedom seeking citizens and security officials noticed this recent event. From the very beginning, the motivation behind this research was never fueled by the academic contribution. It is only a naive hope that a little insight for future security measures is provided against a multitude of threats that look…just like you and me.

VII. Future Work

There are numerous launch points for future research that stem from this initial effort. From this attempt to combine common algorithms to tackle recognition problems utilizing the evolving technology of hyperspectral imagery and video. In the military domain alone, there is an explosion of surveillance capability that is eagerly sought after throughout the Department of Defense.

Contained within recent news accounts on current military operations is a glimpse of the number of growing challenges surrounding the collection of multi-sensor imagery. Earlier this year, ten of the heavily armed and frequently used MQ-9 Reapers, crucial to the success of the war on terror, started to be equipped with a wide-area airborne surveillance sensor called Gorgon Stare. This sensor is capable of providing motion imagery for a four-kilometer radius, day or night, from up to 12 different angles [116]. This sensor will not replace but supplement the current multi-spectral targeting pod on the Reaper that provides full motion video [116]. One of the current operational missions of the Reaper is to hover over buildings looking for insurgents on the move during engagements [116]. From this current mission alone, the benefits are obvious for the tactical impact of developing enhancements for this academic research effort.

But this capability and the vast amount of data that will flow from these unmanned airborne assets is not fully comprehended until the extent of future acquisition

efforts and strategic implications for these assets are taken into account. In 2006, the predecessors of Gorgon Stare were employed by the U.S. Army's Constant Hawk and a year later the U. S. Marine Corps employed there own system Angle Fire [116]. The Gorgon Stare sensor is now being evaluated for placement on RQ-4 Global Hawk, MQ-1 Predator, and Army MQ-1C Sky Warrior [116].

The recent forecast by Secretary of Defense Robert Gates, stated that the military plans to triple its unmanned airborne inventory over the next decade [117]. In addition to this short-term forecast, consider the fact that the Defense Advanced Research Projects Agency (DARPA) is working on a project to increase the total number of views from twelve up to 60 [116]. Recently retired, Deputy Chief of Staff for Intelligence, Surveillance and Reconnaissance (ISR), General David Deptula, stated that the eventual goal will be to study 30 to 60 targets simultaneously with one sensor pod [116]. The increase in sensor capability multiplied by the number of employed airborne assets should give you a feel for the size of the challenge to simply process this data. Next we will turn to the Chief of Staff of the Air Force (CSAF), General Norton Schwartz, for an answer or confirmation of this challenge.

In a recent interview with the CSAF [118], the reporter asked what the impact was for next generation drones considering it takes approximately 70 man-hours to process imagery data for each hour that a Predator is aloft and by his calculations some of these enhanced capabilities will drive that to an astounding 800 man-hours for one hour of flight data [118]. The CSAF's answer was not reassuring, when he simply stated, "the bottom-line is, again, that we cannot continue to throw people at this. We have to find ways to do this better, less manpower-intensive." Keep in mind; the previous examples

and this interview were only focused on air breathing assets. In the last year, the experimental TAC-SAT 3 satellite equipped with a hyperspectral payload was turned over to the Air Force to support operations.

The ability to investigate associated problems to this research has been significantly enhanced by the development of the Multi-Dimensional Face Recognition (MDFR) Graphical User Interface (GUI) and Interactive Development Environment (IDE). Any extension of this research would be well served to explore advanced capabilities within this development platform. Most of the following improvements could be studied using this software tool.

In this research only one illumination setting was explored, that being the brightest lighting that employed three 600 watt halogen bulbs. The reason for this, illustrated earlier was due to experimental nature of the camera and the difficulties with the imagery obtained. Other lighting variations that utilized one of the lights, either from the left, right or center would provide variability that would provide the employment of an algorithm like the previously discussed Fisherface algorithm that has shown its strength in this arena with grayscale images. The exploration and extension of this algorithm would be a tailor made arena to discover and exploit the algorithms contribution to hyperspectral imagery.

The low illumination conditions presented with these single light settings provides an ideal setting that simulates one of the more difficult recognition environments and frankly one that has not been thoroughly explored with the sole exception of thermal sensors and imagery for tracking purposes. These are just some of the opportunities that

naturally grow from the foundation of this research and are available using the existing CMU database.

Although there was considerable experimentation with a combination of agents and algorithm weightings, an opportunity exists to apply an optimization method to the selection of these methods and weightings. Recalling the earlier findings of Jarudi [6], an approach that considers image quality or resolution should be implemented to guide the relative weightings of features for maximum peformance. The work of Friend [123] looked at various information theory-based methods to identify objects most likely to misidentified as well as developing an optimization scheme for selecting appropriate thresholds for classification or label accuracy.

Multispectral segmentation and subsequent fusing has also provided advantages in the recognition of sub features such as the eyes as shown in Boyce's [53] iris recognition research. Unique spectral characteristics also present themselves when observing the human eye. The iris contains melanin with remittance properties that can map to the soft biometric of eye color using multispectral information [53]. With the increasing ability to obtain this information at greater distances, this feature set could easily be integrated into the existing hierarchy. This additional capability could once again span features from soft biometrics, such as eye color, to the very specific, utilizing the complex textural pattern of the iris.

Using a dataset that is not closed, that is all probes are also located in the gallery and vice versa, will change the complexity of the problem and can be used to test the robustness the approach design. Previous pattern recognition research by Friend and Bauer [123] looked at information theory based methods to identify subjects or objects

that were most likely to be misidentified. In Friend's original research [124], objects that were not like any used in the training set, our gallery, are labeled as out of library (OOL) and those subjects that are indistinguishable from more than one subject are subsequently categorized as a non-declaration (NDEC). In these research efforts, the authors looked at OOL and NDEC methodologies that would optimize performance in these more challenging operating conditions. These types of considerations would allow the final design to be better tailored to user requirements and the desired level of effectiveness or robustness.

In a few places in this document, the CSU face recognition system was discussed and used as both a guide and a benchmark for this development. One of the useful elements explored with the capability that is offered by that system was experimentation on the selection of the best distance measurement. As pointed out by some of those findings, the answer to this question is often dependent on the algorithm used and imagery that being processed. The impact of these decisions is not explored in this research and could provide both a needed understanding and perhaps improved performance.

A number of classifiers were employed in this fusion architecture but much like the weightings, the selection of the best set of classifiers is not offered. The work of Chawla [33], [34] on ensemble construction and testing offers some ideas on how to start this investigation. Also important in this effort will be the inquiry into the contribution based on diversity measurements mentioned by Kuncheva [86]. All of these areas, initially applied to face recognition present another opportunity as it is applied to the new and exciting application of hyperspectral face recognition.

The structure of this algorithm allows the input of additional images to fine tune any early matching solution. This happens simultaneously while any number of adjustments can be made with established feedback loops that reduce the gallery size of potential candidates for matching, selection and adjustment of algorithms weightings or even the algorithms themselves. Parallel processing explored in the testing of a range of setting for algorithms would provide additional computational power to enable multiple instantiations of this method for a combination of fusion methods that provide feedback to each other. These methods may seem to be a little excessive given the results already obtained with the basic methodology. The extension of this problem to an uncontrolled environment and the movement to hyperspectral video could easily adjust the problem space to a paradigm that may make this more understandable.

As soon as think about this problem as it is applied to motion video sensors like those being employed in the Gorgon Stare sensor package on an unmanned airborne system, the slow jumpy version of the full motion video feed provides a problem that would require jitter detection [122] and corrective methods for image registration and robust recognition capability.

Although discussed and illustrated with initial testing, the area of disguise and occluded faces in hyperspectral imagery was not included in the performance testing. There is an opportunity to apply and extend the insights and methods provided by the research of Ramanathan [65], Chowdhury [66], and Pavlidis [54]. Certainly, the application of simple spectral angle matching can be implemented to identify subjects that may be trying to conceal their identity with excessive makeup or prosthetics. More complex combinations of algorithms and adaptive processes will ultimately become

useful when applied to these multifaceted problems. The appearance of many of the distinctive features of the ear shown in the hyperspectral image indicates an opportunity to exploit the single modality of hyperspectral imagery and apply this capability to occluded faces. To date this area has not been explored and based on the complementary nature of ear and face features a recognition application that combines the opportunity for front and profile viewpoint features offers an opportunity for the design of a robust identification system.

Performance optimization efforts and adaptive classification system should be studied for use in commercial or operational applications. For these purposes, comparison testing should be conducted against other leading systems and the benchmark of human recognition. Insights provided from cognitive research could be mimicked in this methodology. For instance, the research of Jarudi [6] that illustrated the importance of external features versus internal features was dependent on the distance or quality of the image. This testing plan to incorporate similar findings should look at recognition performance over changes in time, data sample sizes, and viewpoints. A more thorough testing regime can easily be designed that exceeds those initially accomplished in recent Face Recognition Vendor Testing [15]. However, the testing conducted by NIST has set the foundation to for future research to build upon, challenge and bypass the incredible recognition capability of human beings.

Bibliography

[1] Jain, Anil, K., "Biometric Recognition: How Do I Know Who You Are?," *Signal Processing and Communications Applications Conference*, 3-5 (2004).

[2] Jain, Anil, Arun Ross, Salil Prabhakar, "An introduction to biometric recognition", *IEEE Transactions on Circuits and Systems for Video Technology*, 4 - 20 (January 2004).

[3] Jain, A. K., A. Ross and S. Prabhakar, "An Introduction to Biometric Recognition," IEEE Transactions on Circuits and Systems for Video Technology", *Special Issue on Image- and Video-Based Biometrics*, 14.1, 4-20 (January 2004).

[4] Sinha, Pawan, Benjamin Balas, Yuri Ostrovsky, Richard Russell, "Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About," *Proceedings of the IEEE*, 94.11: 1948-1962 (November 2006).

[5] Balas, Benjamin and Pawan Sinha, "Portraits and Perception: Configural Information in Creating and Recognizing face Images," Spatial Vison, 1-17 (2007).

[6] Jarudi, Izzat, Pawan Sinha, "Recognizing degraded faces: Contribution of internal and external features," Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Artificial Intelligence Memo 2003-004 and Center for Biological and Computational Learning Memo 225 (2005).

[7] Valentine, T, "Face-Space Models of Face Recognition," To appear in, Wenger, M. J. & Townsend, J. T. (Eds.) Computational, geometric, and process perspectives on facial cognition: Contexts and challenges. Hillsdale, New Jersey: Lawrence Erlbaum (1999).

[8] Johnston R. A., A. B. Milne, C. Williams, J. Hosie, "Do distinctive faces come from outer space? An investigation of the status of a multidimensional face-space" *Visual Cognition*, 4:59-67 (1997).

[9] Doddington, George R., Liggett, W., Martin, A., Przybocki, M., Reynolds, D., "Sheep, Goats, Lambs, and Wolves, An Analysis of Individual Differences in Speaker Recognition Performance," ICSLP '98 Conference, Sidney, 1-30 (November 1998).

[10] Ellis, H., M. Lewis, "Capgras delusion: a window on face recognition," *Trends in Cognitive Sciences*, 5.4: 149-156 (April 2001).

[11] Tranel, D., A. Damasio, H. Damasio, "Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity," *Neurology*, 690-696 (1988).

[12] Yin, R. K, "Looking at Upside-down Faces," *Journal of Experimental Psychology*, 81.1: 141-145 (1969).

[13] Thompson, Peter, "Margaret Thatcher: A new illusion," *Perception*, 9.4: 483-484 (1980)

[14] Clement, Gilles, *Fundamentals of Space Medicine.* Dordrecht :Kluwer Academic Publishers, 2003.

[15] Phillips, J. P., W Scruggs, A. O'Toole, P. Flynn, K. Bowyer, C. Schott, M. Sharpe, "FRVT 2006 and ICE 2006 Large-Scale Results," 1-55 (Mar 2007).

[16] Ross A. and A. K. Jain, "Multimodal Biometrics: An Overview", *Proceedings of 12th European Signal Processing Conference*, 1221-1224 (September 2004).

[17] Singh, Saurabh, G. Aglika, G. Bebis, I. Pavlidis, "Infrared and Visible Image Fusion for Face Recognition," Computer Vision Laboratory, University of Nevada, Reno, 1-12.

[18] Kong, Seong, J. Heo, B. Abidi, J. Paik, and M. Abidi, "Recent advances in visual and infrared face recognition – a review," *Computer Vision and Image Understanding*, 97: 103-135 (2005).

[19] Kirby, M. and L. Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12.1: 103-108 (January 1990).

[20] Turk, Matthew and Alex Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, 3.1: 71-86 (Winter 1991).

[21] Robinson, Marshall, Matthew Escarra, Jon Krueger, Doug Kochelek, "Face Recognition using Eigenfaces," Houston: Connexions, 2008, 24 October 2008, http://cnx.org/content/col10254/1.2/.

[22] Belhumeur, Peter N., Joao P. Hespanha, and David J. Kriefgman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *1996 European Conference on Computer Vision*, 1-15 (1996).

[23] Duda, Richard, P. Hart, D. Stork, *Pattern Classification*, New York: Wiley, 2001.

[24] Pentland, Alex, Baback Moghaddam, Thad Starner, “View-Based and Modular Eigenspaces for Face Recognition,” M.I.T. Media Laboratory Perceptual Computing Section Technical Report No. 245, 1-7, (1994).

[25] Prince, Simon J. D., James H. Elder, Jonathan Warrell, Fatima M. Felisberti, "Tied Factor Analysis for Face Recognition across Large Pose Differences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30.6: 970-984 (June 2008).

[26] Pan, Zhengjun and Bolouri, H., “High Speed Face Recognition Based on Discrete Cosine Transforms and Neural Networks,” Science and Technology Centre, University of Hertfordshire, 1-12, September 1999.

[27] Er, Meng Joo, Shiqian Wu, Juwei Lu, HockLye Toh, "Face Recognition with Radial Basis Function (RBF) Neural Networks," *IEEE Transactions on Neural Networks*, 13.3: 697-710 (2002).

[28] Hansen, Lars K. and P. Salamon, "Neural Network Ensembles," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12.10: 993-1001 (October 1990).

[29] Abate, Andrea F., M. Nappi, D. Riccio, G. Sabitino, “2D and 3D face recognition: A survey,” *Pattern Recognition Letters*, 28:1885-1906 (January 2007).

[30] Zou, Xuan, J. Kittler, K. Messer, “Illumination Invariant Face Recognition: A Survey,” *Proceedings of IEEE Conference of Biometrics: Theory Applications and Systems*, 1-8 (2007).

[31] Zhao, W., R. Chellappa, P. J. Phillips, A. Rosenfeld, “Face Recognition: A Literature Survey,” *ACM Computing Surveys*, 35.4: 399-458 (December 2003).

[32] Phillips, P. Jonathon, P. Grother, R. Michaels, D. Blackburn, E. Tabassi, M. Bone, “Face Recognition Vendor Test 2002,” Overview and Summary, 1-12 (March (2003).

[33] Chawla, Nitesh V., Kevin W. Bowyer, “Random Subspaces And Subsampling For 2-D Face Recognition,” *CVPR*, San Diego II: 582-589 (June 2005).

[34] Chawla, Nitesh V., Kevin W. Bowyer, “Actively Exploring Creation of Face Space(s) for Improved Face Recognition,” *Association for the Advancement of Artificial Intelligence*, 1-6 (2007).

[35] Penev, Penio S., Lawrence Sirovich, “The Global Dimensionality of Face Space,” *Proceedings of International Conference on Automatic Face and Gesture Recognition, 2000*, 264-270 (2000).

[36] Beveridge, R., Bolme, D., Teixerira, M., and Draper, B., 2003, The CSU Face Identification Evaluation System User's Guide: Version 5.0, Computer Science Department, Colorado State University.

[37] Miller, Philip and Jamie Lyle, "The Effect of Distance Measures on the Recognition Rates of PCA and LDA Based Facial Recognition," *Digital Image Processing*, Clemson University.

[38] Jain, Anil, Sarat Dass, Karthik Nandakumar, "Can Soft Biometric Traits Assist User Recognition," *Proceedings of SPIE*, 5404: 561-572 (2004).

[39] Balci, Koray and Volkan Atalay, "PCA for Gender Estimation: Which Eigenvectors Contribute?," *Proceedings of Sixteenth International Conference on Pattern Recognition, IEEE*, 363-366 (2002).

[40] Moghaddam, Baback and Ming-Hsuan Yang, "Learning Gender with Support Faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24: 707–711 (May 2002).

[41] Phillips, P. Jonathon, H. Moon, S. Rizvi, and P. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:1090–1104 (October 2000).

[42] Kwon, Young Ho and Niels da Vitoria Lobo, "Age Classification from Facial Images," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 762–767 (April 1994).

[43] Horng, Wen-Bing, Cheng-Ping Lee, Chun-Wen Chen, "Classification of Age Groups Based on Facial Features," *Tamkang Journal of Science and Engineering*, 4.3: 183-192 (2001).

[44] Mukaida, Shigeru and Hiroshi Ando, "Extraction and Manipulation of Wrinkles and Spots for Facial Image Synthesis," *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, 1-6 (2004).

[45] Gutta, Srinvis, Harry Wecheler, "Gender and Ethnic Classification of Human Faces Using Hybrid Classifiers," *Proc. Int. Joint Conference on Neural Networks*, 6: 4084-4089 (1999).

[46] Marcel, Sebastien, and Samy Bengio, "Improving Face Verification Using Skin Color Information," *International Conference on Pattern Recognition*, 1-4 (2002).

[47] Sun, Q., and M. Fairchild, "Statistical Characterization of Spectral Reflectances and Its Application to Human Portraiture Spectral Estimation," *The Journal of Imaging Science and Technology*, 46.6: 498-506 (2002).

[48] Anderson, R., J. Parrish, "The Optics of Human Skin," *The Journal of Investigative Dermatology*, 77:13-19 (1981).

[49] Storring, M., H. Andersen, E. Granum, "Physics Based Modelling of Human Skin Colour Under Mixed Illuminants," *Robotics and Autonomous Systems*, 35: 131-142, (2001).

[50] Jablonski, N., G. Chaplin, "The Evolution of Human Skin Coloration," *Journal of Human Evolution*, 39: 57-106 (2000).

[51] Daugman, J., "Face and Gesture Recognition: Overview," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19.7: 675-676 (July 1997).

[52] Bataille, V., "Genetics of Risk Factors for Melanoma: An Adult Twin Study of Nevi and Freckles," *Journal of the National Cancer Institute*, 92.6: 457-463 (March 2000).

[53] Boyce, Christopher, Arun Ross, Matthew Monaco, Lawrence Hornak, Xin Li, "Multispectral Iris Analysis: A Preliminary Study," *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop*, 1-9 (2006).

[54] Pavlidis, Ioannis, Peter Symosek, "The Imaging Issue in an Automatic Face/Disguise Detection System," *Proceedings IEEE Workshop on Computer Vision beyond the Visible Spectrum: Methods and Applications*, 15-24 (2000).

[55] Nunez, Abel, "Modeling the Reflectance of Human Skin," 1-16 (Draft 2008).

[56] Jenkins, R. and A.M. Burton, "100% Accuracy in Automatic Face Recognition," *Science*, 319: (January 2008).

[57] Er, Meng Joo, W. Chen, Shiqian Wu, "High Speed Face Recognition Based on Discrete Cosine Transform and RBF Neural Networks." *IEEE Transactions on Neural Networks*, 16.3: 679-91 (2005).

[58] Pan, Zhihong, Glenn Healey, Manish Prasad, Bruce Tromberg, "Hyperspectral Face Recognition under Variable Outdoor Illumination," *Proceedings of SPIE*, 5425:520-529 (2004).

[59] FieldSpec® remote cosine receptor (RCR), Analytical Spectral Devices, Inc. (1999). http://www.asdi.com/FR-Foreoptics.pdf

[60] Lowe, David G., "Object Recognition from Local Scale-Invariant Features," *Proceedings of the International Conference on Computer Vision*, Corfu, 1-8 (September 1999).

[61] Lowe, David G., “Distinctive Image Features from Scale-Invariant Keypoints,” *International Journal of Computer Vision*, 1-28 (2004).

[62] Estrada, F., A. Jensen, D. Fleet, “Local Features Tutorial,” (November 2004).

[63] Mian, Ajmal, Mohammed Bennamoun, Robyn Owens, “An Efficient Multimodal 2D-3D Hybrid Approach to Automatic Face Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29.11: 1927-1943 (November 2007).

[64] Luo, Jun, Yong Ma, Erina Takikawa, Shihong Lao, Masato Kawade, Bao-Liang Lu, "Person-Specific SIFT Features for Face Recognition," *IEEE International Conference on Acoustics, Speech and Signal Processing, 2007.*, 2: 593-596 (April 2007).

[65] Ramanathan, Narayanan and Rama Chellapa, “Facial Similarity Across Age, Disguise, Illumination and Pose,” *Proceedings of International Conference on Image Processing*, 3: 1999-2002 (2004).

[66] Chowdhury,Roy., R. Chellappa, N. Ramanathan, and N. Koterba, “About face,” *National Geographic*, 18–19 (November 2003).

[67] Singh, Richa, Mayank Vasta, Afzel Noore, “Face Recognition with Disguise and Single Gallery Images,” *Image and Vision Computing*, 1-13 (July 2007).

[68] Nunez, Abel, M. J. Mendenhall, “Remote Sensing of Human Skin and its Chromophores in the Near Infrared,” *IEEE Transactions on Geoscience and Remote Sensing*, 1-11 (Received September 2008).

[69] Paquit, Vincent, J. Price, F. M´eriaudeau, K. Tobin, T. Ferrell, "Combining near-infrared illuminants to optimize venous imaging," *Proceedings of the SPIE Medical Imaging 2007*, 1-9 (February 2007).

[70] Yan, Ping and K. Bowyer, “Biometric Recognition Using 3D Ear Shape,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29.8: 1-12 (August 2007).

[71] Chen, Hui and B. Bhanu, “Human Ear Recogntion,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29.4: 718-737 (April 2007).

[72] Pavlidis, Ioannis, James Levine, Paulette Baukol, "Thermal Imaging for Anxiety Detection," *Proceedings of IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications, 2000*, 104-109 (2000).

[73] Jain A. K. and A. Ross, "Multibiometric Systems", *Communications of the ACM, Special Issue on Multimodal Interfaces*, 47.1: 34-40 (January 2004).

[74] Singh, Richa, Mayank Vasta, Afzel Noore, “Hierarchical Fusion of Multi Spectral Face Images for Improved Recognition Performance,” *Information Fusion*, 9.2: 200-210 (April 2008).

[75] Pan, Zhihong, Glenn Healey, Bruce Tromberg, “Multiband and Spectral Eigenfaces for Face Recognition in Hyperspectral Images” *Proceedings of SPIE*, 5779:144-151 (2005).

[76] Roli, F., “Fusion 2002 Tutorial, Fusion of Multiple Classifiers”

[77] Korves, H., L. Nadel, B. Ulery, “Multi-biometric Fusion-From Research to Operations,” *Mitretek Systems*, Presentation (June 2005).

[78] Indovina, M., U. Uludag, R. Snelick, A. Mink and A. Jain, "Multimodal Biometric Authentication Methods: A COTS Approach", *Proceedings of Workshop on Multimodal User Authentication*, 99-106 (December 2003).

[79] Ross, A. and A. Jain , “Information Fusion in Biometrics,” *Pattern Recognition Letters*, 2115-2125 (September 2003).

[80] Jain, Anil, Karthik Nandakumar, Xiaoguang Lu, Unsang Park, “Integrating Faces Fingerprints, and Soft Biometric Traits for User Recognition,” *Proceedings of Biometric Authentication Workshop*, 259-269 (May 2004).

[81] Schubert, C., M. Oxley, K. Bauer, “Determining the ROC Curve of Fused Independent Classification Systems”, To be submitted to Elsevier Science, 1-20 (2008 Draft).

[82] Wang, R., B. Bhanu, "Performance Prediction for Multimodal Biometrics," *18th International Conference on Pattern Recognition*, 3: 586-589 (2006).

[83] Viswanathan, R. and P.K. Varshney, “Distributed Detection with Multiple Sensors, Part I: Fundamentals,” *Proceedings of IEEE*, 54-63 (January 1997).

[84] Veeramachaneni, K., L. A. Osadciw, P. K. Varshney, "An Adaptive Multimodal Biometric Management Algorithm," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on Systems, Man and Cybernetics*, 35.3: 344-356 (August 2005).

[85] Chair, Z., P. Varshney, “Optimal Data Fusion in Multiple Sensor Detection Systems,” *IEEE Transactions on Aerospace and Electronic Systems*, AES-22.1 (January 1986).

[86] Kuncheva, Ludmila, *Combining Pattern Classifiers, Methods and Algorithms*, Wiley and Sons, New Jersey, 2004.

[87] Surowiecki, James, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies Societies and Nations*, Double Day, 2004.

[88] Sensor Reading Group, "What is Information Theory?, The Basics," Presentation, University of Cal Tech, October 2003. http://robotics.caltech.edu/readinggroup/2003-Oct-17.pdf

[89] Bowyer, Kevin, Chang, K., Flynn, P., Chen, X., "Face Recognition Using 2-D, 3-D, and Infrared: Is Multimodal Better Than Multisample?," *Proceedings of the IEEE*, 94.11: 2000-2012 (November 2006).

[90] Socolinsky, Diego A. and Andrea Selinger, "Thermal Face Recognition in an Operational Scenario," *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1-8 (2004).

[91] Dowdall, Jonathan, Ioannis Pavlidis, George Bebis, "Face Detection in the Near-IR Spectrum, *Image and Vision Computing*, 21:7, 565–578 (2003.).

[92] Pan, Zhihong, Glenn Healey, Manish Prasad, Bruce Tromberg, "Hyperspectral Face Recognition in for Homeland Security," *Proceedings of SPIE*, 5074: 767-776 (2003).

[93] Robila, Stephan, "Toward Hyperspectral Face Recognition", *Proceedings of SPIE – IS&T Electronic Imaging*, 6812: 1-9 (2008).

[94] Robila, S. A., "Using Spectral Distances for Speedup in Hyperspectral Image Processing", *International Journal of Remote Sensing*, 26.24: 5629-5650 (2005).

[95] Chou, Yi-Ting and P.Bajcsy, "Toward Face Detection, Pose Estimation and Human Recognition from Hyperspectral Imagery," *Technical Report NCSA-ALG04-0005 (*October 2004).

[96] Pan, Zhihong, Glenn Healey, Manish Prasad, Bruce Tromberg, "Face Recognition in Hyperspectral Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25.12: 1552-1560 (December 2003).

[97] Elbakary, M. I., M. S. Alam, M. S. Aslan, "Face Recognition Algorithm in Hyperspectral Imagery by Employing the K-means Method and the Mahalanobis Distance," *Proceedings of SPIE*, 6697:1-9 (2007).

[98] Denes, Louis J., Peter Metes, and Yanxi. Liu, "Hyperspectral Face Database," *Tech. Report CMU-RI-TR-02-25*, Robotics Institute, Carnegie Mellon University (October 2002).

[99] http://www.equinoxsensors.com/products/HID.html

[100] http://www.ee.oulu.fi/research/imag/color/color.php?page=facedatabase/facedatabase.php

[101] Rogers, Steven K., Kabrisky, M., Bauer, K., Oxley, M., "Computing Machinery and Intelligence Amplification," *Computational Intelligence: The Experts Speak*, 3: 25-44, 2003

[102] Rogers, Steven K., Types of Qualia – Version 15, Presentation, November 2009.

[103] Rogers, S., M. Kabrisky, K. Bauer, M. Oxley, A. Rogers, *QUEST: QUalia Exploitation of Sensor Technology* (Draft 2008).

[104] Varshney, P. K., "Multisensor Data Fusion," *Electronics and Communication Engineering Journal*, 9.6: 245-253 (December 1997).

[105] Li, Stan and Jain, A. K. Jain. *Handbook of Face Recognition*. US: Springer, 2005.

[106] Sakhi, Omid Bonakdar, Face Detection Matlab code Version : 5.0, Feed Forward Neural Network Code and Demonstration Images, May 2007.

[107] Bar, Moshe and M. Neta, "The proactive brain: Using rudimentary information to make predictive judgements," *Journal of Consumer Behavior*, 7: 319-330 (2008).

[108] Bar, Moshe, "The proactive brain: using analogies and associations to generate predictions," *TRENDS in Cognitive Sciences*, 11.7:280-9 (June 2007).

[109] Ando, Hiroshi, "ATR Human Information Processing Laboratories," International Media Technology Workshop on Abstract Perception, Japan (January 1994).

[110] Samal, Ashok and P. Iyengar, Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey, *Pattern Recognition*, 25.1:65-77 (1992).

[111] Hofstadter, Douglas, *Godel, Escher, Bach: An Eternal Golden Braid*, Basic Books, New York, 1979.

[112] Gonzalez, R. C., Woods, R. E., and Eddins, S. L., Digital Imaging Processing Using MATLAB, Prentice Hall, Upper Saddle River, NJ, 2004.

[113] Billor, Nedret, Hadi, A., Velleman, P., "BACON: blocked adaptive computationally efficient outlier nominators," *Computational Statistics and Data Analysis*, 34: 279-298 (2000).

[114] McGill, Robert, Tukey, J., Larsen, W., "Variations of Box Plots," The American Statistician, 32.1:12-16 (February 1978).

[115]"Box Plot – MATLAB." *MathWorks – MATLAB and Simulink for Technical Computing*. Web. 07 Nov 10. http://www.mathworks.com/help/toolbox/stats/boxplot.html.

[116] Hoffman, Michael, "New Reaper sensors offer bigger picture," Air Force Times, 19 Feb 2009. http://www.airforcetimes.com/news/2009/02/airforce_WAAS_021609/

[117] Capaccio, Tony, "Boeing Co. Drones Play Pivotal Role In War On Taliban, Al-Qaeda," BusinessWeek.com, 19 Jul 2010. http://ebird.osd.mil/ebfiles/e201007217648441.html

[118] Muradian, Vago, "Interview: Gen Norton Schwartz, U. S. Air Force Chief Of Staff," Defense News, 10 May 2010. http://ebird.osd.mil/ebfiles/e20100510751136.html

[119] AT&T the Database of Faces, AT&T Laboratories Cambridge, http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

[120] Rogers, Steve, QUEST: Integrated Human computer 'solutions' for the 'wicked' problems we face, Presentation, April 2010.

[121] Smetek, Timothy and K. Bauer, A Comparison of Multivariate Outlier Detection Methods for Finding Hyperspectral Anomalies, 74th MORS Symposium, 1-33, 2006.

[122] Simonson, Katherine M., Ma, T. J., Robust Real-Time Change Detection in High Jitter, Sandia Report SAND2009-5546, 1-41, September 2009.

[123] Friend, Mark A, Bauer, Kenneth W., "An Entropy-based Scheme for Automatic Target Recognition," *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 1-12, January 2010.

[124] Friend, Mark A. *Combat identification with synthetic aperture radar, out-of-library identification, and non-declarations*. Air Force Institute of Technology (AU), Wright Patterson AFB OH, May 2008 (AAT 3288578).

[125] Wilkes, Malini. "Disguise Raises Airline Security Concern." Liveshots. FOX News Network, 07 Nov 2010. Web. 07 Nov 2010. http://liveshots.blogs.foxnews.com/2010/11/07/disguise-raises-airline-security-concern/

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*		3. DATES COVERED *(From – To)*
14-03-2011	Dissertation	Sept 2006 –Mar 2011

4. TITLE AND SUBTITLE

QUEST HIERARCHY FOR HYPERSPECTRAL FACE RECOGNITION

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Ryer, David M., Lt Col, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Street, Building 642
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/DS/ENS/10-03M

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFRL/RY
2241 Avionics Circle
Area B Bldg 620
WPAFB OH 45433-7765

10. SPONSOR/MONITOR'S ACRONYM(S)
NASIC/DAIA AFRL/RYZT

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Face recognition is an attractive biometric due to the ease in which photographs of the human face can be acquired and processed. The non-intrusive ability of many surveillance systems permits face recognition applications to be used in a myriad of environments. Despite decades of impressive research in this area, face recognition still struggles with variations in illumination, pose and expression not to mention the larger challenge of willful circumvention. The integration of supporting contextual information in a fusion hierarchy known as QUalia Exploitation of Sensor Technology (QUEST) is a novel approach for hyperspectral face recognition that results in performance advantages and a robustness not seen in leading face recognition methodologies.

This research demonstrates a method for the exploitation of hyperspectral imagery and the intelligent processing of contextual layers of spatial, spectral, and temporal information. This approach illustrates the benefit of integrating spatial and spectral domains of imagery for the automatic extraction and integration of novel soft features (biometric). The establishment of the QUEST methodology for face recognition results in an engineering advantage in both performance and efficiency compared to leading and classical face recognition techniques. An interactive environment for the testing and expansion of this recognition framework is also provided.

15. SUBJECT TERMS

Face Recognition, Hyperspectral Imagery, Biometrics, QUEST, Image Processing, Pattern Recognition

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Kenneth W. Bauer (ENS)
U	U	U	UU	228	19b. TELEPHONE NUMBER *(Include area code)* (937) 255-6565, ext 4328; e-mail: Kenneth.Bauer@afit.edu